

University of Dundee

DOCTOR OF MEDICINE

Evaluation of SELDI-TOF MS as a tool in colorectal cancer screening

Henderson, Nikola Alexandra

Award date:
2014

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

DOCTOR OF MEDICINE

Evaluation of SELDI-TOF MS as a tool in colorectal cancer screening

Nikola Alexandra Henderson

2014

University of Dundee

Conditions for Use and Duplication

Copyright of this work belongs to the author unless otherwise identified in the body of the thesis. It is permitted to use and duplicate this work only for personal and non-commercial research, study or criticism/review. You must obtain prior written consent from the author for any other use. Any quotation from this thesis must be acknowledged using the normal academic conventions. It is not permitted to supply the whole or part of this thesis to any other person or to post the same on any website or other online location without the prior written consent of the author. Contact the Discovery team (discovery@dundee.ac.uk) with any queries about the use or acknowledgement of this work.

Evaluation of SELDI-TOF MS as a tool
in colorectal cancer screening.

Nikola Alexandra Henderson

in submission for the degree of Doctor of Medicine

2014

Contents

Chapter 1 Introduction

Part 1: Colorectal cancer

1.1.1.1	Colorectal cancer	Page 10
1.1.1.2	Pathogenesis of colorectal cancer	Page 15
1.1.1.3	Management, survival and outcome	Page 17
1.1.2.1	Screening for Colorectal Cancer	Page 19
1.1.2.2	Stool Based Testing	Page 19
1.1.2.3	Faecal Immunochemical Testing	page 20
1.1.2.4	Wilson and Junger Criteria	Page 22
1.1.2.5	CRC as a candidate for screening	Page 22
1.1.3.1	Stool based testing	Page 24
1.1.3.2	Leaked Markers	Page 28
1.1.3.3	Secreted Markers	Page 31
1.1.3.4	Exfoliated Markers	Page 31
1.1.4.1	Endoscopic Screening	Page 33
1.1.4.2	Flexible sigmoidoscopy	Page 33
1.1.4.3	Colonoscopy	Page 36
1.1.4.4	CTC Colonography	Page 36
1.1.5.	Serum markers	Page 39
1.1.6	Conclusions	Page 42

Part 2: Clinical Applications of SELDI-TOF Proteomics

1.2.1	Novel Technologies	Page 43
1.2.2.1	SELDI-TOF MS	Page 45
1.2.3.1	Proteomics in Diagnosis	Page 49
1.2.3.2	Ovarian Cancer	Page 50
1.2.3.3	Head and Neck Cancer	Page 50
1.2.3.4	Prostate Cancer	Page 51
1.2.3.5	Pancreatic Cancer	Page 52
1.2.3.6	Breast Cancer	Page 52
1.2.3.7	Colorectal Cancer	Page 53
1.2.4.1	Other Screening Approaches	Page 54
1.2.4.2	Cell Lines	Page 54
1.2.4.3	Cervical Cancer and Cytology	Page 55
1.2.4.4	Breast Cancer and Nipple Aspirate	Page 55

	1.2.4.5	Renal Cancer and Urine	Page 55
	1.2.4.6	Monitoring Treatment Response	Page 56
	1.2.4.7	An Update on SELDI-TOF	Page 57
	1.2.5.	Conclusions	Page 57
	1.3	Aims of Study	Page 61
Chapter 2	Materials and Methods		
	Part 1: Clinical Methods		
	2.1.1	Patient Selection	Page 62
	2.1.2	Collection and Storage of Serum	Page 63
	2.1.3	Collection and Storage of Stool	Page 64
	Part 2: Laboratory Methods		
	2.2.1.1	Solutions and Buffers for SELDI	Page 66
	2.2.1.2	Solutions and Buffers for Tumour	Page 69
	2.2.1.3	Optimisation of SELDI-TOF MS	Page 70
	2.2.1.4	Preparation of the Arrays	Page 71
	2.2.2.1	Tumour analysis	Page 74
	2.2.2.2	Protein extraction using SIGMA	Page 74
	2.2.3.	Laser Capture Microdissection	Page 74
	2.2.4	Depletion of serum	Page 77
	2.2.5	Miscellaneous experiments	Page 78
Chapter 3	Analysis of Colorectal Cancer and Normal Mucosa		
	3.1	Introduction	Page 80
	3.2	Optimisation of method	Page 81
	3.3	Analysis of tumours IMAC30 chip	Page 84
	3.4	Laser Capture Microdissection	Page 92
Chapter 4	Analysis of Human Serum		
	4.1	Introduction	Page 93
	4.2	Results	Page 95
	4.3	Serum analysis	Page 96

4.4	Analysis of 10 cancer serum samples compared to 10 normal serum samples on the H50 chip optimised for the low molecular weight region	Page 98
4.5	Analysis of 92 controls compared to 92 cancers using the H50 chip	Page 104
4.6	Analysis of serum on the H50 ProteinChip following serum depletion	Page 108
4.7	Comparison of peaks in serum with those found in tumour	Page 111

Chapter 5 Analysis of Human Faeces

5.1	Introduction	Page 112
5.2	Results	Page 113
5.3	SELDI-TOF MS of Stool	Page 114
5.3.1	Comparison of cancer and controls	Page 116
5.3.2	Comparison of FOB negative and positive	Page 118
5.3.3	Comparison of FOB cancer and FOB negative normal	Page 120
5.3.4	Comparison of FOB negative control and FOB positive normal	Page 121
5.3.5	FOB positive cancer & FOB negative cancer	Page 121
5.3.6	FOB negative cancer & FOB negative normal	Page 122
5.3.7	FOB negative normal & FOB positive normal	Page 122
5.3.8	Degradation over time	Page 123
5.4	Comparison of peaks in stool with those found in tumour	Page 125

Chapter 6 Discussion

6.1	Study design and sample collection	Page 126
6.2	SELDI TOF MS	Page 132
6.3	Analysis of Tumours	Page 134
6.4	Analysis of Serum	Page 135
6.5	Analysis of Faeces	Page 137
6.6	Overall comparison of peaks	Page 139

Conclusions and Future Directions	Page 140
-----------------------------------	----------

References	Page 143
------------	----------

I, Nikola Henderson declare that I am the author of this thesis; that, unless otherwise stated, all references have been consulted by myself; that the work of which this thesis is a record has been done by me, and that it has not been previously accepted for a higher degree.

I wish to thank my supervisor Professor Bob Steele for all his advice, his understanding and his support over the years. Being his research fellow was a great privilege and a joy.

Abstract

Aim: To assess SELDI-TOF MS technology as a tool for biomarker discovery in the stool and serum of colorectal cancer patients.

Materials and Methods:

1. Initially a technique of analysis was developed and optimised using tumour samples and matched normal mucosa, obtained from the Tayside Tissue Bank. These samples were then analysed using SELDI on a PBS II Protein Chip Mass Spectrometer to identify abundant proteins.
2. A technique of stool preparation and subsequent SELDI analysis was developed and then optimised (CM10 chip at pH4) to allow comparison of faecal samples from cancer and controls. Faecal samples were then collected from cancer patients and controls and analysed. In addition, FOB testing was carried out on all stool samples from cancer and controls and subgroup analysis of spectras was performed controlling for FOB status.
3. A test set of cancer and normal serum samples was used to optimise the method of analysis using 4 different chip surfaces at differing pH. Serum samples were collected from cancer patients and normal controls and were analysed on the H50 chip. Serum was then depleted of major proteins in an attempt to improve the detection of peaks. The mass spectra from each sample type were compared to identify any common protein peaks.

Results:

1. Tumour analysis methods were optimised using an initial 4 samples of tumour and normal mucosa. Analysis of 8 further paired samples showed protein peaks at 2826, 3374, 3444, 3489 and 10854 Da which were abundant in tumour and reduced in the normal mucosa.

2. In serum analysis the initial experiment of 10 cancer versus 10 normal revealed 4 peaks on the H50 chip (3479, 3364, 3434, 3700 Da) that had significantly higher mass to charge ratios in cancer. The experiment was repeated on the H50 chip using 92 cancers and 92 controls and 5 different peaks were identified (7901, 8124, 8566, 8799 and 17409Da) as being significant but these had higher mass to charge ratios in the controls. After depletion of the serum samples of albumin, transferrin, haptoglobin, anti-trypsin, IgG and IgA SELDI-TOF analysis showed a greatly reduced profile that yielded no meaningful spectra.

3. Stool analysis revealed 5 protein peaks (4633, 16511, 33423, 37087 and 47026 Da) in colorectal cancer patients, which were absent in stools from controls with a sensitivity of 83% when using all 5 peaks. Degradation of the spectra was observed after prolonged storage of stool samples.

Conclusions: A method of stool analysis has been developed that yielded valid peaks differentiating between cancer and normal, which warrant further research through protein identification. Serum analysis was not reproducible across experiments and depletion of major proteins failed to reveal the sub-proteome raising doubts about whether discovery-based serum proteomics can accurately detect cancer. SELDI-TOF was not able to demonstrate that any of the peaks present in the tumour analysis were present in the stool or the serum samples.

ABBREVIATIONS

%	Percent
2D	Two dimensional
2D PAGE	Two dimensional gel electrophoresis
3D	Three dimensional
APC	adenomatous polyposis coli
BDH	maker of chaps
C	degrees Celsius
CA	California
CEA	Carcinoembryonic antigen
CHAPS	Cell extract buffer
CHCA	a-cyano-4-hydroxycinnamic acid
CRC	Colorectal cancer
CT	Computer assisted Tomogram
CTC	Computer assisted Tomogram Colonography
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme Linked Immuno Assay
FAP	Familial adenomatous polyposis
gFOBt	Guaiac Faecal occult blood test
FOBt	Faecal occult blood test
G	G force
H ₂ O	Water
HEPES	(4(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNPCC	Hereditary non polypotic colon cancer
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IN	Indeannapolis
iTRAQ	Isobaric tag for relative and absolute quantification
KCl	Potassium chloride
LCM	Laser capture microdissection
M	Molar
MALDI-TOF	Matrix assisted laser desorption time of flight mass spectrometry
mM	Millimolar
MSI	Microsatellite instability
MUC	Mucins
Na	Sodium
Na ₂ HPO ₄	Sodium hydrogen phosphate
NaCl	Sodium chloride
NaOH	Sodium hydroxide

NHS	National health service
NSAID	Non steroidal anti-inflammatory drug
PBS	Phosphate buffered saline
pH	pH
PSA	Prostate specific antigen
RCC	Renal cell carcinoma
RCT	Randomised controlled trial
RNA	Ribonucleic acid
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS PAGE	SDS Gel electrophoresis
SELDI-TOF MS	Surface enhanced laser desorption time of flight mass spectrometry
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TME	Total mesorectal excision
TNM	Tumour Node Metastases
TOF	Time of flight
TRIS-HCl	Tri(hydroxymethyl)amino methane
UICC	International
UK	United Kingdom
uL	Microlitres
USA	United States of America

CHAPTER 1

INTRODUCTION

Part 1 Colorectal cancer

1.1.1.1 Colorectal cancer

Colorectal cancer is the third most common cancer diagnosed in the United Kingdom and accounts for 10% of all cancer deaths(Mayor 1998). 35 300 cases were diagnosed in England and Wales in 2000, accounting for 13% of all cancers diagnosed that year and 842 men and 713 women died of the disease in Scotland in 2002(Statistics 2005). Incidence is higher in males and there is also considerable variation amongst different ethnic groups. It is a disease of the developed world and despite an increasing risk in many countries, overall survival rates are improving(Scholefield, Moss et al. 2002). Survival correlates with stage and ranges from 95% five year survival for Dukes' A disease, 80% Dukes' B and 45% for Dukes' C cancer, so that the benefits of diagnosing colorectal cancers early are easily appreciated(Coleman 2004).

Sporadic colorectal cancers are more common than familial cases and account for almost 90% of diagnosed tumours(Hardy, Meltzer et al. 2000). Genetic and environmental factors play an important role and although many of these lifestyle and environmental factors especially those associated with lifestyle can be controlled by the individual, there is also a strong association with previous colorectal polyps and a familial association that is not accounted for by the same genetic mechanism as the hereditary colorectal cancers. Familial adenomatous polyposis (FAP), previously known as polyposis coli and part of the spectrum of genetic changes that constitute Gardner's Syndrome (FAP with extra colonic manifestations) is caused by mutations in the APC gene. This autosomal dominant disease leads to the development of hundreds of adenomas in the colon and if left untreated will

lead to the development of colorectal cancer by middle age in almost all patients. Removal of the colorectum does not entirely protect the patients from the mortality and morbidity of this condition as duodenal tumours and intra-abdominal desmoids can develop as extra-colonic manifestations(Hardy, Meltzer et al. 2000).

The other well known hereditary colorectal cancer is hereditary non-polyposis colorectal cancer (HNPCC), another autosomal dominant condition with incomplete penetrance, also known as the Lynch syndrome. The modified Amsterdam criteria (Table 1) are used to provide the diagnosis of Lynch syndrome as the phenotype is not diagnostic as with FAP. The original Amsterdam criteria only included colonic cancers but it is now known that endometrial and small bowel adenocarcinomas occur as part of this syndrome.

Modified Amsterdam Criteria	
<ul style="list-style-type: none"> • Three or more cases of colorectal cancer in a minimum of two generations • One affected individual must be a first degree relative of the other two (or more) cases • Colorectal cancer can be replaced by endometrial or small bowel adenocarcinoma • Familial adenomatous polyposis must be excluded 	

Table 1.1 Modified Amsterdam Criteria(Llor, Pons et al. 2005)

Not all colorectal cancer families fulfill the Amsterdam Criteria and some other syndromes such as Turcot's, Peutz-Jeghers or juvenile polyposis will present with a higher risk of colorectal cancer than normal. More common are inflammatory bowel disease patients (Crohn's disease and ulcerative colitis) who are at higher risk of developing a malignancy; in ulcerative colitis colorectal cancer accounts for a third of deaths(Gyde 1990). The extent of inflammation and disease duration are the main risk factors(Cairns and Scholefield 2002).

Table 1.2 shows the risk factors associated with colorectal cancer.

<p>Sporadic colorectal cancer</p> <p>Old age Male sex Previous cholecystectomy (colonic only) Ureterocolic anastomosis Nulliparity, late first pregnancy, early menopause</p> <p><i>Environmental factors:</i> Meat and fat rich diet Obesity Smoking Low fibre, folate and calcium diet Sedentary lifestyle High alcohol consumption Previous irradiation Occupational hazards (asbestos exposure)</p> <p><i>Personal history of tumours:</i> Colorectal polyps Previous colorectal cancer Previous small bowel, ovarian, breast or endometrial cancer</p> <p>Familial colorectal cancer One affected first-degree relative increases risk 2.3 fold Two or more affected first degree relatives increases risk 4.25 fold If less <45 years old increases risk 3.9 fold Familial history of colorectal adenoma increases risk 2 fold</p> <p>Colorectal cancer in inflammatory bowel disease Ulcerative colitis Crohn's disease</p> <p>Hereditary colorectal cancer Hereditary non-polypoid colorectal cancer (HNPCC) <i>Polypoid syndromes:</i> familial adenomatous polyposis (FAP), Gardner's syndrome, Turcot's syndrome, flat adenoma syndrome <i>Hamartomatous polyposis syndromes</i> Peutz-Jeghers syndrome, Cowden syndrome</p>
--

Table 1.2. Factors associated with the varying typed of colorectal cancer.

Dukes' original 1932 classification of rectal tumours has been extrapolated to include all tumours of the colorectum (Dukes 1932). There are three original stages, A, B and C, C is further subdivided to represent the level of nodal disease and a more advanced stage D has been added on to include cancers that have metastasised.

Dukes' A tumour confined to bowel

Dukes' B breached bowel wall, no node disease

Dukes' C1 Nodes positive, highest lymph node clear

Dukes' C2 Nodes positive, highest lymph node diseased

Figure 1.1 shows this in diagrammatic form.

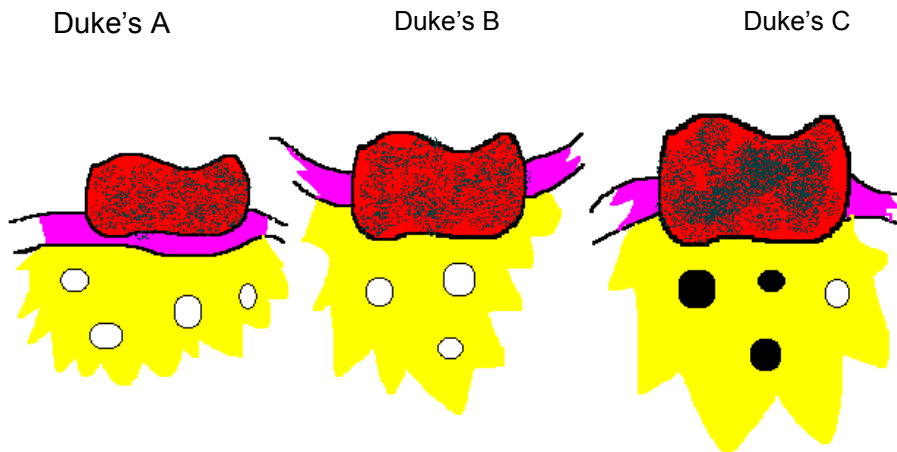


Figure 1.1

Representation of Dukes' Staging. The tumour is shown in red, bowel wall in pink and the nodes are the white (negative) and black (positive) spherical shapes within the yellow fat.

Further treatment as well as prognosis is guided by the Dukes' stage, additionally it is now common practice to use the TNM stage for each case. TNM is the acknowledged method used around the world to stage malignant tumours; it is promoted by the UICC (International Union Against Cancer) as the gold standard in cancer staging. TNM staging is based on the size of the primary tumour, the T value, the presence or absence of disease in the regional lymph nodes, the N value and the presence of distant metastases, the M value. Other parameters can be included in the TNM, for example G for grade of tumour and V for invasion into vessels although none of these are mandatory for staging. Dukes' stage can be derived from the TNM stage and both are used interchangeably in clinical practice. Table 1.3 shows the TNM stages for colorectal cancer.

T stage	Primary Tumour		
TX	Cannot be assessed		
T0	Carcinoma in situ		
T1	Submucosal invasion		
T2	Muscularis propria invasion		
T3	Through muscularis propria into subserosa		
T4	Tumour invades other organs/structures/visceral peritoneum		
N stage	Regional lymph nodes		
NX	No regional lymph node metastases		
N0	1 to 3 nodes positive		
N1	4 or more nodes positive		
N2			
M stage	Distant metastases		
MX	Cannot be assessed		
M0	None		
M1	Present		
	T	N	M
Stage I	T1, T2	N0	M0
Stage IIA	T3	N0	M0
Stage IIB	T4	N0	M0
Stage IIIA	T1,T2	N1	M0
Stage IIIB	T3,T4	N1	M0
Stage IIIC	Any T	N2	M0
Stage IV	Any T	Any N	M1

Table 1.3

UICC TNM staging for colorectal cancer. Version 7 2009

Pathogenesis: The adenoma carcinoma sequence

1.1.1.2 Pathogenesis of CRC

Most colorectal cancers arise as part of the adenoma-carcinoma sequence (Morson, 1978)

There is extensive epidemiological, clinicopathological and genetic evidence for this sequence which is shown schematically below. It is thought that normal epithelium, in response to environmental factors and dietary carcinogens causes a progression through low grade dysplasia to higher grade dysplasia and then eventual carcinoma (Leslie, Carey et al. 2002). The evidence for the colorectal adenoma carcinoma sequence is convincing; oncogenes, tumour suppressor genes and DNA repair genes all contribute to the stepwise progression to cancer. The APC tumour suppressor gene, located on chromosome 5q21 is mutated early in the sequence and mutation of this gene results in an intracellular accumulation of β -catenin, a key protein in cell transcription (Rustgi 1993). APC mutations are present in a large percentage of colorectal cancers and also in adenomas.

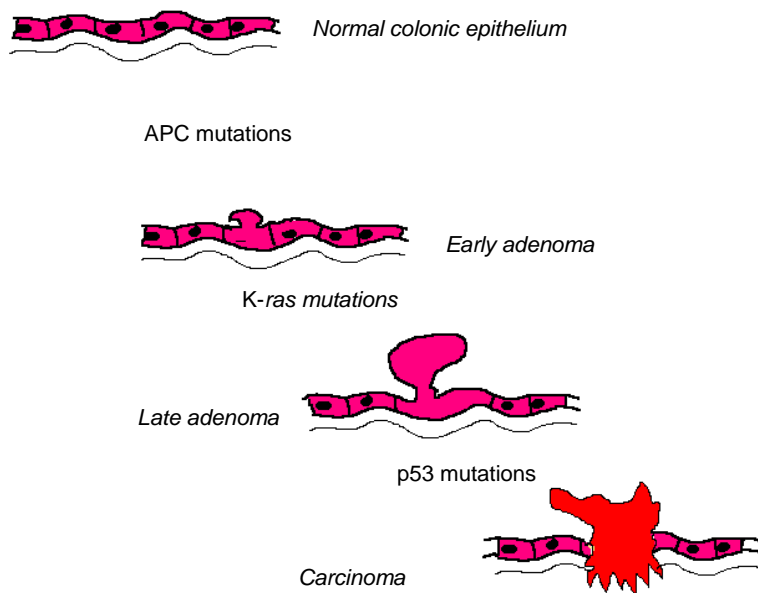


Figure 1.2 Adenoma-Carcinoma Sequence
Normal colonic epithelium progresses to adenoma and then cancer through cumulative genetic changes

Following on from APC mutations are the Kras and p53 alterations(Kemp, Thirlwell et al. 2004). K ras mutations occur commonly in adenomas and carcinomas and are part of the stepwise progression from normal mucosa through to invasive cancer. The p53 gene located on the long arm of chromosome 17 is lost in many human cancers including colorectal cancers; accumulation of p53 protein within the cell may cause the progression of adenoma into carcinoma. Microsatellite instability (MSI) is a further category of abnormality that is present in adenomas as well as cancers(Calvert and Frucht 2002). However, what is clear from the genetics of colorectal cancer, is that this is a vastly complex area where genetic mutations do not occur in isolation and a number of further mechanisms, microsatellite instability and mismatch repair proteins, 18q loss and the other alterations in methylation status lead to progression to cancer. Not all adenomas become cancers, those at high risk of malignant transformation are the large, sessile or flat polyps, those with a severely dysplastic or villous architecture and presence of multiple other polyps(Hardy, Meltzer et al. 2000). Polyps are generally considered low risk if they are small, mildly dysplastic and pedunculated. The genetics of colorectal cancer must be applicable in the clinical arena and this possibility is becoming increasingly likely with time.

Not all adenomas become cancers, those at high risk of malignant transformation are the large, sessile or flat adenomas, those with a severely dysplastic or villous architecture and presence of multiple other polyps(Hardy, Meltzer et al. 2000). Adenomas are generally considered low risk if they are small, mildly dysplastic and pedunculated. The genetics of

colorectal cancer must be applicable in the clinical arena and this possibility is becoming increasingly likely.

1.1.1.3 Management, survival and outcome

Following staging of the patients, treatment is usually planned by a multidisciplinary team of surgeons, radiologists, pathologists and specialist nursing staff and is based on clinical and radiological staging. A CT scan of the liver and lungs excludes or reveals metastases and this also provides information on the extent of tumour invasion and lymphadenopathy. For local staging in rectal cancer MRI is now preferred owing to its enhanced ability to distinguish tissue planes (Brown and Daniels 2005). The whole colon must be visualised by colonoscopy and ideally this is undertaken prior to surgery or very soon after as 5% of patients will have a synchronous tumour (Finan, Ritchie et al. 1987). The treatment of colonic cancer is aimed at cure through surgical resection with the hemicolectomy forming the basis of treatment. In rectal cancer, Heald's total mesorectal excision (TME) has reduced morbidity from pelvic nerve damage and local recurrence rates as well as improving survival (Heald and Ryall 1986). Laparoscopic surgery has been an option in colonic surgery for some time but as yet it has not achieved the same degree of acceptance of laparoscopic cholecystectomy. There are a number of reasons for this but evidence is now supporting laparoscopic hemicolectomy in cancer as being as good a treatment as open surgery and in some series it has been shown to be better (Lacy, Garcia-Valdecasas et al. 1995). In rectal cancer it has also been shown to be safe and effective (Dulucq, Wintringer et al. 2005). Adjuvant therapies consist of radio or chemotherapy or a combination of the two. In rectal cancer radiotherapy is used neoadjuvantly and occasionally postoperatively to prevent local recurrence. In Dukes' C cancers, adjuvant chemotherapy is given if the patient has a good performance status and adequate organ function, the most recent guidelines advocate combination

chemotherapy in those that can tolerate it (Scottish Intercollegiate Guidelines Network 2011).

1.1.2.1 Screening for Colorectal Cancer

Screening for cancer in the United Kingdom is well established for both cervical cancer and breast cancer and these programmes are very successful at detecting early disease and preventing cancer deaths (Blanks, Moss et al. 2000). In contrast to this, screening for colorectal cancer is in its infancy despite being the second most common cause of cancer deaths and there being 35,000 new cases diagnosed in the UK every year (UK 2006). The lifetime risk of developing colorectal cancer in Scotland is 5.5% for males and 4.4% for females (Executive 2007). Colorectal cancer is highly curable when diagnosed at an early, asymptomatic stage but when not detected until later stages it is aggressively fatal (Woodman, Prior et al. 1995; Gatta, Faivre et al. 1998).

1.1.2.2 Stool Based Testing

In Scotland, the government has rolled out a national screening programme using a guaiac based Faecal Occult Blood Test (gFOBT) for all males and females registered with a general practitioner between 50 and 74 years of age. This commenced in the spring of 2007 and after a staged roll out was nationwide by 2009 (Steele, McClements et al. 2009). The programme is based on biannual gFOB testing that if positive results in an invitation to attend for a colonoscopy. FOBT based screening programmes lead to a reduction in mortality from colorectal cancer as demonstrated by population based randomised controlled trials (Hewitson, Glasziou et al. 2008) and by analysis of pilot studies in Scotland (Libby, Brewster et al.). FOB testing comprises sending a small stool sample to a central testing

laboratory where the sample is tested for the presence of haemoglobin. The most widely used FOBT works on the principle that guaiac, made from the resin of the *Guajacum* tree, is turned blue by pseudoperoxidases from haemoglobin, in the presence of hydrogen peroxide (Anker, Christensen et al. 1974). There are different FOB tests on the market and the most commonly used, Haemoccult is a guaiac based test that detects the presence of haemoglobin in the stool including animal haemoglobin but also reacts to peroxidases unrelated to haemoglobin (Scholefield 2000). The unrehydrated guaiac method has a sensitivity of around 50% in a screening context, (Fraser, Matthew et al. 2006). The Minnesota trial, one of four large randomised controlled trials that showed a reduction in mortality, found that re-hydration raised the sensitivity for cancers from 80.8% to 92.2%, with a converse decrease in specificity from 97.7% to 90.4% giving a net reduction in the positive predictive value (Mandel, Bond et al. 1993). The mortality reduction in the four large trials of FOB screening is shown in Table 1.4. These important trials from Minnesota, Nottingham, Denmark and Sweden showed a reduction in colorectal cancer mortality with FOB testing ranging from 15% to 33%. Additionally, the Minnesota trial also supported the adenoma carcinoma sequence in that it demonstrated a reduction in cancer incidence over time, presumably due to therapeutic polypectomy. The other trials did not show this reduction in incidence presumably because they were associated with a lower positivity and thus a lower colonoscopy rate. A Cochrane review in the form of meta-analysis of these trials has been carried out showing a reduction of 15% overall from FOB testing. (Hewitson, Glasziou et al. 2008)

Lead Investigator, date , journal)	Result	PPV of Haemoccult for CRC
Mandel et al(Mandel, Bond et al. 1993) NEJM 1993 “Minnesota Trial”	33% reduction in CRC mortality over 13 years (in annual screening) 21% in biennial	2.2% rehydrated 5.6% unrehydrated
Hardcastle et al(Hardcastle, Chamberlain et al. 1996) Lancet 1996 “Nottingham Trial”	15% reduction in mortality over 7.8 years (biennial screening)	9.9% 1 st screen 11.9% 2 nd screen
Kronborg et al(Kronborg, Fenger et al. 1996) Lancet 1996 “Danish trial”	18% reduction in mortality at 10 years. (biennial screening)	17.7% 1 st screen 8.4% 2 nd screen 16.3% 3 rd screen 10.8% 4 th screen 10.2% 5 th screen
Lindholm et al(Lindholm, Brevinge et al. 2008) British Journal of Surgery “Swedish trial”	16% reduction in mortality at 19 years.	5.2% overall for cancer 14% overall for large adenomas

Table 1.4 Summary of the results of the four large FOB trials that showed a reduction in mortality

1.1.2.3 Faecal Immunochemical Testing

The other forms of FOB testing are the haemoporphyrin tests, such as Haemaquant®, which react to porphyrins from haemoglobin and give an estimate of the amount of blood, and the immunochemical tests, such as HemeSelect®, which identify only intact human haemoglobin. Immunochemical faecal occult haemoglobin testing, known as iFOBT or FIT testing, offers advantages over the gFOBT. An antibody is used to detect intact globin

protein from human haemoglobin, thus permitting no cross reaction with animal derived blood protein or with plant peroxidases as can occur with gFOB testing. No dietary restriction is necessary when using the FIT test. Additionally, as globin from the upper GI tract is readily digested in the gut, blood derived from the upper GI tract either from pathological or iatrogenic reasons (eg aspirin therapy) does not cause a positive test. Fraser et al have shown that FIT testing in the gFOBT positive population can reduce the number of false positive tests and lead to fewer screening colonoscopies in normal patients (Fraser, Matthew et al. 2006). They have also demonstrated that the median concentration of haemoglobin present in the stool varies in normal patients (13.5ng/ml), high risk adenomatous disease (65.6ng/ml) and in cancer (165ng/ml), with wide standard deviations (Digby, Fraser et al.). Immunochemical FOBT testing is more sensitive for cancers than for benign neoplasia and is as good at detecting proximal as distal disease (Levi, 2007).

Systematic review and meta-analysis of 19 studies utilising iFOBT (from 1996 to 2013) concluded that the test has high overall diagnostic accuracy for CRC detection, with a variable degree of success depending on the cutoff value chosen to represent a positive result, sensitivity of 79% with a specificity of 94% was found on this analysis (Lee, 2014). Reducing the cut off point for a positive can lead to an increase in the detection of adenomatous disease but does not show a corresponding increase in detection of cancer. There is not as yet a randomised controlled trial using an iFOBT in a screening context.

1.1.2.4 Wilson and Junger Criteria

In order to successfully implement a screening program the criteria proposed by Wilson and Junger must be met (Wilson JM 1968).

- The condition should be an important health problem
- There should be an accepted effective treatment for patients with recognised disease
- Facilities for further diagnosis and treatment should be available
- There should be a recognisable latent or early symptomatic stage
- There should be a suitable test or examination
- The test should be acceptable to the population
- The natural history of the condition, including development from latent to declared disease should be adequately understood
- There should be an agreed policy on whom to treat as patients
- The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole
- Case finding should be a continuous project and not a “once and for all” project.

1.1.2.5 CRC as a candidate for screening

Colorectal cancer would seem to be an ideal candidate for screening on the basis of these criteria. Colorectal cancer is common and is increasing in prevalence in countries that are becoming “Westernised”(Sung, Lau et al. 2005). In addition, there is good evidence that cancers develop from adenomas(Baba 1997) and that removal of these can reduce cancer incidence(Atkin, Edwards et al.; Mandel, Church et al. 2000). Thus screening has the potential to reduce cancer incidence provided adenomas can be detected. Treatment of

colorectal cancer is well established and curative surgery forms the mainstay of management with survival correlating to clinical stage. Surgical management continues to advance with new techniques being introduced that may reduce mortality and morbidity even further(Lacy, Garcia-Valdecasas et al. 1995),(Dulucq, Wintringer et al. 2005),(Heald and Ryall 1986). Tumours that previously would have seemed inoperable are now treatable with adjuvant therapies; radiotherapy, chemotherapy and both in combination can be used and there are new chemotherapeutic treatments that use less toxic oral drugs rather than the older infusion regimes(Weitz, Koch et al.; Kang, Chang et al. 2005).

Following a positive FOB test facilities are available for further management in the form of a colonoscopy and although this is an area of concern for the already overburdened endoscopy service in the NHS(Bowles, Leicester et al. 2004) the UK governments Departments of Health have given an undertaking that the needs of screen detected individuals will be met. Cost is an important consideration and FOBT screening has been shown to be both cost-effective(Frazier, Colditz et al. 2000) and sufficiently acceptable to the screening population(Homa, Brzosko et al. 2005). Thus colorectal cancer is a convincing candidate disease for a screening programme.

Determining the screening population, the method of screening and the screening interval is slightly less clear cut. There are agreed guidelines for those in high risk groups (familial adenomatous polyposis, Lynch syndrome and previous CRC patients) and these are generally not included in the discussion for mass population studies(Cairns and Scholefield 2002). Only 7% of colorectal cancers occur below the age of 50 so most screening is commenced from this age onwards(Imperiale, Wagner et al. 2002). The pilot study in the UK screened only those up to the age of 69 but the colorectal cancer screening programme in Scotland now includes those up to the age of 74 years as 50% of all colorectal cancers occur in those

over 70 years old; debate about cost of screening and number of life years saved by intervening in this elderly group is ongoing(Walter, Lewis et al. 2005).

There are a variety of different screening modalities from the high cost, highly invasive colonoscopy based screening programme to the low cost non-invasive stool based tests and there is debate over the best method. As outlined above it was shown by the Minnesota study in 1993 that FOB based screening could reduce mortality from colorectal cancer(Mandel, Bond et al. 1993) and this was confirmed when both the Nottingham and the Danish studies were published in 1996(Hardcastle, Chamberlain et al. 1996; Kronborg, Fenger et al. 1996), and the more recent Swedish study(Lindholm, Brevinge et al. 2008). (Cochrane meta-analysis(Hewitson, Glasziou et al. 2008).) In spite of this, many years later the debate continues about other modalities demonstrating the difficulties in deciding how best to screen.

1.1.3.1 Stool based testing

There are two main groups of colorectal cancer screening tests, visualisation of the colorectum either directly using endoscopes or indirectly using radiological investigations and stool based testing, looking for haemoglobin or other markers of neoplasia. The majority of studies into colorectal cancer screening have evaluated the use of FOB tests, flexible sigmoidoscopy or colonoscopy and the advantages and disadvantages of each are addressed in table 1.5. This table shows that there is relatively little cost and little inconvenience in performing an FOB test, including the cost to society and to the patient. In contrast a colonoscopy requires at least one and usually two days off work, sedation which results in an inability to drive home afterwards, all of which cause a relatively large amount of inconvenience.

	FOB	SIGMOIDOSCOPY	COLONOSCOPY	CT COLONOGRAPHY
Setting	Home	Endoscopy suite	Endoscopy suite	Radiology department
Bowel prep	None	Minimal	Total bowel prep	Total bowel prep
Fasting	None	None	24-48 hours	24 hours
Level of training required	None	Minimal	Highly skilled	Highly skilled
Duration	n/a	20 minutes	Up to 40 minutes	20 minutes
Cost	Cheap	moderate	expensive	Expensive
Sensitivity for cancers	70%	95%	100% in best hands “Gold Standard”	
Day off work	No	No	At least one day	
Morbidity	None	Low	Moderate	Low
Extent of colon seen/tested	Entire colorectum	Up to 60cm	Entire colorectum	Entire colorectum
Further investigation needed	Yes if positive	Yes if polyp/cancer	No	Yes for tissue diagnosis
Post procedure care	None	None	Transport home Supervision	None

Table 1.5 Comparison of screening modalities. Adapted from Macafee et al (Macafee and Scholefield 2003)

Different health care systems and societies accept different levels of cost, sensitivity and specificity and risk of complications. Where health care is delivered by the private sector and screening is not largely opportunistic there is a preference towards more accurate and more expensive forms of screening. The first round of the UK wide pilot showed that of the 478,250 people invited to take part, 56.8% returned a completed FOB test. The overall rate of a positive test result was 1.9% and the rate for detecting cancer was 1.62 per 1000 people screened with a positive predictive value of 10.9% for cancer and 35% for adenoma (UK Colorectal Cancer Screening Pilot Group 2004).

The stage of screen detected cancers was also in contrast to the symptomatic stage distribution of cancers, with many more early and potentially curable cancers in the screen detected group; 48% of all screen detected cancers were Dukes' stage A, and 1% were Dukes' D at time of diagnosis (UK Colorectal Cancer Screening Pilot Group 2004). This is in marked contrast to the symptomatic patients and this is illustrated in Figure 1.3.

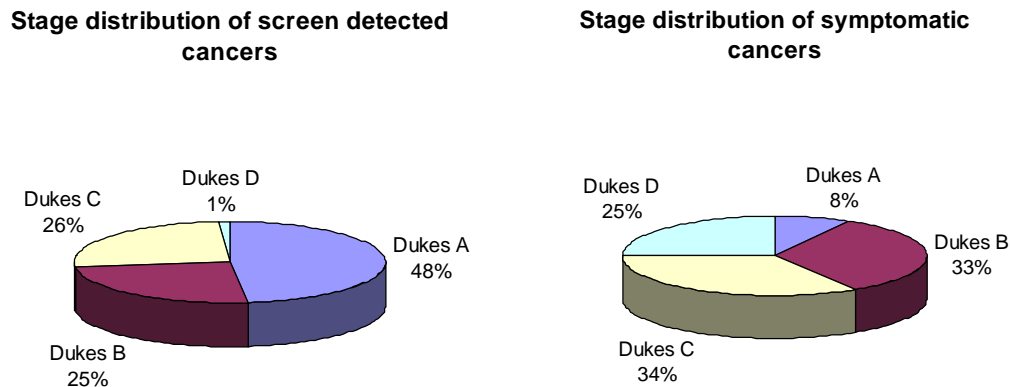


Figure 1.3 Comparison by Dukes's stage of screen detected versus symptomatic cancers

The 57% return rate was lower in men particularly in the 50 to 55 age group. Those who live in deprived areas and in areas where there is a large proportion of people from the Indian subcontinent are also less likely to return a sample (Greiner, Engelman et al. 2004).

Stool based testing is the initial method of detection in many CRC screening programmes and investigation of a positive test can involve either direct visualisation of the colon using colonoscopy or sigmoidoscopy or indirect radiological methods such as barium enema (usually combined with a sigmoidoscopy) or CT colonography.

Stool is easily accessible on an almost daily basis, it is transportable and involves no painful or cathartic procedure for retrieval and as it journeys through the colon it reflects the surface of the entire colorectum. The already established method of screening, the FOBT can be performed on mailed in specimens of stool, therefore allowing widespread access to screening across all geographical locations.

However the specificity and sensitivity of the FOBT is far less than the gold standard investigation, colonoscopy (Villa, Dugani et al. 1996). The inherent limitations using blood for screening is that occult bleeding is intermittent when it occurs from early neoplasms and in addition, it commonly occurs in a large number of benign conditions of the colorectum (Ahlquist, McGill et al. 1989). No matter how sensitive the FOB test, there is no way of avoiding the fact that on the days of testing the tumour may not be bleeding. To allow for this the participants in the Scottish screening programmes are asked to take a sample from three separate bowel movements. If a strong positive result is returned an invite to attend for colonoscopy is made, in the event of a weak positive then further stool testing is offered in the form of a more specific immunochemical Hema-Screen test, if this is negative then the participant is returned to the asymptomatic group to await further screening in 2 years and if positive they are invited for colonoscopy.

Patterns of bleeding from occult colorectal cancers have been studied and are known to be highly variable; in particular, the absence of symptoms of colorectal cancer has been shown to correlate with a normal faecal haemoglobin measurement. (Doran and Hardcastle 1982;

Ahlquist, McGill et al. 1989) In some centres, to avoid the problem of contamination of endogenous and animal haemoglobin participants are asked to abstain from red meat and NSAID's prior to testing, though it has been shown that dietary restriction is a barrier to screening (Cole and Young 2001). Screening for colorectal cancer has caused controversy with many differing opinions on the best way to proceed and despite the evidence, critics of FOB based screening advocate caution in rolling out national programmes (Towler, Irwig et al. 1998)

An ideal stool marker would of course be present in high concentrations in early disease, be stable in stool and transport media, be easily assayed and be 100% specific and 100% sensitive. Ahlquist has proposed a classification system for stool tumour markers based on their mechanism of luminal entry; leaked, secreted or exfoliated (Ahlquist and Gilbert 1996).

1.1.3.2 Leaked Markers

Haemoglobin from blood is the most common leaked marker, other markers that are leaked from tumour into stool have been detected and explored as potential screening tests though none of these have been adopted as a screening tool. Levels of calprotectin, a leukocyte-derived protein in the cytosol of neutrophil granulocytes, are increased in the stools of colorectal cancer patients (Roseth, Fagerhol et al. 1992). Studies have shown that the levels of haemoglobin and of calprotectin are not linear so it is likely not to be leaked into stool in the same way as blood (Gilbert, Ahlquist et al. 1996). It is thought that calprotectin gains luminal access via interstitial leukocyte migration, which could be a less variable mechanism of entry than bleeding (Gilbert, Ahlquist et al. 1996). Testing for calprotectin has been shown to be less useful than a sigmoidoscopy as a screening tool and there has not been a large comparison of FOB versus calprotectin despite initial hopes that it may replace FOB

testing(Roseth, Kristinsson et al. 1993; Hoff, Grotmol et al. 2004) Calprotectin detection is facilitated by its stability in stool and reliable estimates can be obtained in samples of only 5 g. Research in Scandinavia has revealed that 79% of patients with CRC had a raised stool calprotectin level but the specificity of this as a marker has not been fully evaluated. It is known to be raised in benign inflammatory conditions as well as neoplasia(Kristinsson, Armbruster et al. 2001). Current evidence suggests that its higher sensitivity in some series compared to the FOB comes with a lower specificity(Tibble, Sigthorsson et al. 2001) which in a prospective trial was found to be 63% (Limburg, Devens et al. 2003).

Lactoferrin is a neutrophil derived glycoprotein that has been measured in a variety of disease states. It was initially described as a marker in inflammatory bowel disease, along with other proteins found in stool(Sudo, Igawa et al. 1993; Sugi, Saitoh et al. 1996). Further evaluation of this glycoprotein has shown it is elevated in Travel Associated Diarrhoea, Clostridium difficile diarrhoea and in paediatric gastrointestinal disease (Scerpella, Okhuysen et al. 1994; Steiner, Flores et al. 1997; Ruiz-Pelaez and Mattar 1999). In addition, a prospective pilot study measuring lactoferrin by enzyme linked immunoassay (ELISA) found that lactoferrin was as useful as FOB in determining cancer from non-cancer in fresh stool samples(Saitoh, Kojima et al. 2000). Recent work has indicated a role for Lactoferrin measurement in post-operative inflammatory bowel disease patients to determine inflammatory diarrhea pouchitis and also in monitoring response to Infliximab therapy(Buderus, Boone et al. 2004; Parsi, Shen et al. 2004).

Lysozyme is an innate, non-immunologic antibacterial enzyme produced by the Paneth cells of the upper intestinal tract. Lysozyme is not normally secreted in the lower intestinal tract but it is by colorectal neoplasias and although unstable in stool, it is easily detectable and quantifiable(Brouwer and Smekens 1991; Dubrow, Kim et al. 1992). In addition to being

present in the stool of patients with cancers of the colorectum, evidence suggests that this marker is present in adenoma tissue, although it has not been demonstrated in the stool samples of patients with adenomas(Rubio 2003).

1.1.3.3 Secreted Markers

It may be that substances secreted from the tumour into the lumen of the bowel may indicate disease. Mucins found in stool are high-molecular weight epithelial glycoproteins with a high content of clustered oligosaccharides. There are two structurally and functionally distinct classes of mucins, the secreted gel-forming mucins (MUC2, MUC5AC, MUC5B, and MUC6)(Byrd and Bresalier 2004) and the transmembrane mucins (MUC1, MUC3A, MUC3B, MUC4, MUC12, MUC17). In addition a number of more difficult to classify mucins have been identified (MUC7, MUC8, MUC9, MUC13, MUC15, MUC16) (Limburg, Ahlquist et al. 2000). MUC1 mucin, as detected immunologically, is increased in expression in colon cancers, but unfortunately does not discriminate accurately between cancers and controls(Limburg, Ahlquist et al. 2000) . Expression of MUC2 secreted gel-forming mucin is generally decreased in colorectal adenocarcinoma, but preserved in mucinous carcinomas, a distinct subtype of colon cancer associated with microsatellite instability(Ajioka, Allison et al. 1996; Song, Byrd et al. 2005).

1.1.3.4 Exfoliated Markers

An estimated 10^{10} colonocytes per day are continuously shed into the lumen of the bowel and eliminated from the body in faecal matter with the entire lining of the colon being renewed every 3 or 4 days, offering a continual supply of marker release in contrast to the varied nature of bleeding(Sidransky, Tokino et al. 1992). Colonocytes are lost from the surface by not only the sloughing caused by the passage of stools but also through apoptosis and subsequent engulfment(Eastwood 1977; Eastwood 1992; Eastwood 1995). One of the main features of malignant transformation is the decrease in apoptosis which leads to an

increase in the amount of cellular material in the stool samples accordingly colorectal cancer patients have been shown to have a significantly higher amount of DNA in their stool samples and a larger number of colonocytes (Villa, Dugani et al. 1996). Despite the relatively small surface area of a neoplasm in contrast to the rest of the colonic mucosal surface, an estimated 24% of recovered DNA in colorectal cancer patients is tumour DNA (Ahlquist, Harrington et al. 2000). In addition to loss of apoptosis there is increased proliferation of cells and less adhesiveness between cells. Due to the large amount of cellular debris within faecal samples and our ability to perform polymerase chain reactions to amplify areas of DNA, genetic tests are increasingly being explored as potential stool based tests for colorectal cancer and K-ras mutations were identified as early as 1992 in stool samples (Sidransky, Tokino et al. 1992). DNA is stable within stool, and fragments of colonocytes are easily assayed. However, whole colonocytes are not stable in stool and isolating them can be challenging (Davies, Freeman et al. 2002). The amount of non-human DNA in stools may also be a confounding factor, and distal tumours seem easier to detect than proximal tumours, perhaps because the environment in the right colon is more hostile to DNA.

The genetic mutations associated with colorectal cancer are heterogeneous and a panel of DNA changes is more useful than one single abnormality. Multi target DNA markers are by far the most promising, covering many known genetic changes in colon cancer, but they are costly. Ahlquist used a panel that gave a sensitivity for cancer of 91% and 82% for adenomas with a specificity of 93%. Excluding K-ras from the panel, sensitivities for cancer were unchanged but decreased slightly for adenomas to 73%, whilst specificity increased to 100% for cancer (Ahlquist, Skoletsky et al. 2000).

DNA tests are very cancer specific and so false positives are not returned from patients with benign disease, unlike FOBT. The tests that have been described in the literature are highly

sensitive and specific and like the FOB specimens can be collected at home and do not require bowel preparation. However, considerable further work is required to decide the most appropriate combination. In addition, stool DNA testing may answer more questions than just the presence or absence of colorectal cancer and other cancers more proximal in the gastrointestinal tract can also be detected using DNA based stool tests(Ahlquist DA 2000).

1.1.4.1 Endoscopic Screening

Endoscopic methods of screening have been proposed using either sigmoidoscopy or colonoscopy; colonoscopy is widely regarded as an impractical method in the UK due to cost, the potential complications (including death) and lack of trained staff to perform the huge number of screening tests that would need to be done. In the United States, where the private sector and insurance companies shoulder the cost, there is a growing body of support that promotes colonoscopy as the primary means of screening and it is recommended by the American College of Gastroenterologists as the preferred mode of screening.(Bhattacharya and Sack 1996; Podolsky 2000) (Rex, Johnson et al. 2000) Many of the studies evaluating endoscopic methods of screening are case control studies where the end point is incidence and not a reduction in mortality(Selby, Friedman et al. 1992; Winawer, Flehinger et al. 1993; Segnan, Senore et al. 2005).

1.1.4.2 Flexible sigmoidoscopy

Flexible sigmoidoscopy (FS) has been proposed as a screening tool as most cancers and polyps occur in the area of the colon that the flexible sigmoidoscope can reach (60cm from anal margin)(Atkin, Cuzick et al. 1993). The finding of a cancer at sigmoidoscopy or of a

“high risk” polyp can be used to trigger a completion colonoscopy to visualise the whole colon to screen for a synchronous pathology. The benefits of flexible sigmoidoscopy over colonoscopy are less training time needed for the endoscopist, less time off work for the patient, less cost and less bowel preparation and a theoretical corresponding higher uptake. The baseline findings of a UK trial which sub selected only motivated patients to be randomised reported an uptake of 71%. Of those screened, 72.3% had no abnormality and of the abnormal group, 4.7% had high risk lesions and 0.3% had cancer(UK Flexible Sigmoidoscopy Screening Trial Investigators, 2002).

The study published in 2010 by Atkin et al involved 14 UK centres and included 170 432 patients from a normal screening population, between the ages of 55 and 64 years. They were offered a once-only flexible sigmoidoscopy (57 237 participants assigned to FS, 71% uptake). In per protocol analysis, colorectal cancer incidence was decreased by 33% overall, by 50% for distal colon and rectal cancers and mortality was reduced by 43%. These results are impressive but the incidence of right sided cancers was not decreased.

The likelihood of having no distal lesion and a synchronous caecal or other right sided cancer was thought to be rare, but a study published in 2000 showed that by failing to fully investigate the colon with colonoscopy and giving those with no distal polyps on sigmoidoscopy a clean bill of health will result in half the cases of advanced proximal neoplasia being missed(Imperiale, Wagner et al. 2000). However, other studies have not supported this finding and a more recent study looked at polyps in the distal colon detected during colonoscopy and assessed risk depending on polyp type. Advanced neoplasia occurred in the right side in just 2% of patients with no polyps, 2% with distal hyperplastic polyps and 4% of those with adenomas. Based on these findings the authors suggested that the discovery of hyperplastic polyps at sigmoidoscopy does not necessitate further

colonoscopy(Lin, Schembre et al. 2005). A meta analysis of this topic, by Dave et al concluded that in an asymptomatic person, a distal hyperplastic polyp is associated with a 21% to 25% risk for any proximal neoplasia and a 4% to 5% risk of advanced proximal neoplasia(Dave, Hui et al. 2003). Meta analysis of studies using the screening population in America has shown that distal adenomatous polyps, regardless of size, are associated with an increased prevalence of synchronous proximal neoplasia with an odds ratio of 2.4.(Lewis, Ng et al. 2003).

In recent years there has been a proximal migration in cancer locations and now caecal cancers and ascending colon cancers account for an increasingly large proportion of all tumours. (Levi, Randimbison et al. 1993). The biology of right sided cancers may be different to those that occur on the left and in the rectum and certainly right sided cancers are likely to occur in older patients(Slater, Papatestas et al. 1982). FS has been unfavourably compared to performing a mammogram on just one breast and may not be acceptable to the population on the grounds that it is only investigating the highest risk area of the colon and is ignoring what may be just around the corner(2002; Macafee and Scholefield 2003). Despite these concerns, flexible sigmoidoscopy is being rolled out in England starting from March 2013 for men and women aged 55 to 60 years, before FOBt screening starts. Another option is to use flexible sigmoidoscopy in direct combination with FOBt. Several studies have detected more pathology using FOBt with FS as than from using the FOB alone. However, the evidence points to poor compliance in the arms randomized to FS and FOB instead of just FOB(Berry, Clarke et al. 1997; Rasmussen, Kronborg et al. 1999).

1.1.4.3 Colonoscopy

Although colonoscopy is unlikely to be adopted by publicly funded health care systems, it is advocated by the American Society of Gastroenterology as the preferred method of screening; a colonoscopy every ten years for those over 50 is recommended (Walsh and Terdiman 2003)

Currently there are no completed randomized controlled trials to support its use in this context but there are a large number of case control studies that have shown a decrease in incidence of colorectal cancer following a colonoscopy. The protective effect of having a normal colonoscopy is thought to last ten years although this is based on evidence derived from studies involving flexible sigmoidoscopy (Selby, Friedman et al. 1992). A cancer incidence of less than one percent over 5 years following a normal colonoscopy has been shown by smaller case control studies, when individuals are in low risk groups and asymptomatic. (Rex 1996; Zaubler 1997)

The risk from a screening colonoscopy also must be considered; in expert hands there is little risk of colonic perforation occurring but there is a morbidity of 0.3% associated with the procedure including a risk of death (Walsh and Terdiman 2003) Cost is one of the most prohibitive factors, it has been estimated that a colonoscopy costs £3000 per life year saved in contrast to just £1000 for an FOBT and £1500 for a flexible sigmoidoscopy (Scholefield 2000)

1.1.4.4. CT Colonography

This relatively novel radiological technique allows excellent visualisation of the adequately prepared colon using multi-slice helical CT scanning which can be used to create a 3-Dimensional image or “virtual colonoscopy”. As in colonoscopy, the patient is again

required to take the bowel cleansing preparations necessary to purge all faecal material from the colon. The colon is inflated with gas and then the patient is scanned in the supine and prone positions, the data being displayed as a 2 or 3-D image.

The advantages of this technique are that the patient does not have to undergo a colonoscopy unless there is pathology that requires treatment or further investigation. It also has been suggested that patients prefer some degree of choice as they may not want a small polyp removed and prefer to be followed up over time by repeat CTC as the fate of such small polyps may not involve malignant progression (Eide 1991). In addition there are potential social benefits; there is no sedation required so that the patient can drive themselves home from the hospital and they do not require supervision afterwards.

There are disadvantages, however. There is no means of performing any therapeutic procedure or of obtaining tissue for diagnosis, discovery of pathology necessitates further colonoscopy with more bowel preparation and time off work and false positives are common as diverticular disease, prominent ileocaecal valves and thick haustral folds can all mimic polyps or cancers. Other problems, particularly common in the elderly screening population, are related to metal artefacts (such as hip prostheses) causing streaking on the CT scan image. In addition, there is the problem of delivering very high doses of radiation over a period of time which carries the albeit small, theoretical risk of causing cancer (Brenner and Georgsson 2005).

The sensitivity of CTC has been found to be highly variable in one study. Comparing the findings of CTC by three different highly trained radiologists with colonoscopy immediately afterwards showed that CTC interpretation detected an average of 63% (variability of 32% to 74%) of the 59 polyps greater than 1 cm seen at colonoscopy (Johnson, Harmsen et al. 2003). The low pick up rate in general as well as the high interobserver variability is

concerning. CTC is also not useful for identifying flat adenomas and small adenomas with sensitivity as low as 13% in one series(Fidler, Johnson et al. 2002). Even with increasingly detailed 3D reconstruction of the segments of interest it is still easy to miss a premalignant lesion less than 1cm, though some would argue that it is not necessary to detect adenomas less than this size(Atkin 2003).

Ethical problems and treatment dilemmas are created when CT scans detect other pathologies that are not related to colorectal cancer screening. The screening population are high risk for other pathologies and two studies in the United States, from the same hospital, have a large number of incidental findings. Hara et al in the Mayo clinic found 151 pathologies in 109 of their 264 patients that were not related to their colon. 23% of these were “highly important” findings which resulted in further investigation (e.g. renal mass, abdominal aortic aneurysm, etc)(Hara, Johnson et al. 2000). A different group in the same department of the Mayo clinic published further results just three years later; they found in a larger sample size they had a similar pick up of a huge number of findings, but that just 10% of them were of high clinical importance(Gluecker, Johnson et al. 2003).

Where CT colonography may be useful is in screening patients when it is not possible to perform a colonoscopy due to frailty or if colonoscopy would require a general anaesthetic. It is useful to diagnose disease even when the intention is not curative as it can guide treatment and allow a prognosis to be given. Current practice is to perform a Barium enema if it is not possible to complete a colonoscopy but this technique does not visualise the sigmoid colon well. There is research ongoing into improving CTC using methods such as electronically subtracting stool from the images by giving the patient faecal tagging preparations, low residue diets and novel software packages. If the technique can be further evolved to remove the need for bowel preparation it may well be that CT colonography will

have a role in screening but it is unlikely to be applicable in mass population studies(Nicholson, Barro et al. 2005). MRI scanning would be ideal as there is no radiation involved and this is subject to ongoing evaluation.

1.1.5 Serum Markers

It can be argued that there is little to be gained from developing tests that detect advanced cancers, what is urgently needed is a biomarker for early disease. The most established marker in colorectal cancer is CEA (carcinoembryonic antigen) and this used in follow-up but not in the diagnosis of primary, curable disease(McArdle 2000). CEA is a high molecular weight glycoprotein that is thought to play a role in cell adhesion and apoptosis and it has been shown that injection of CEA into mice enhances metastases(Hammarstrom 1999). Various studies have looked at CEA sensitivities for each Dukes' stage, and one used CEA in combination with CA 242(Kim, Fernandes et al. 2003) as illustrated in Table 2.

Stage	CEA alone	CA 242 alone	CEA and CA 242
Dukes A	27.8%	11.1%	33.3%
Dukes B	32.4%	16.2%	48.6%
Dukes C	32.1%	30.8%	40.7%
Dukes D	66.7%	50%	72.5%

Table 1.6 Sensitivity of serum biomarker for each Dukes stage comparing CEA alone, CA242 and the two markers combined.

For CEA to be useful in screening it would need to detect early disease which it does not; CEA detection in a normal population would result in a failure to detect almost half of the cancers(Duffy, van Dalen et al. 2003). CEA is neither sensitive nor specific for colorectal cancer with estimates of approximately 87% specificity and 40% sensitivity(Fletcher 1986).

Serum CA 19-9 (Morita, Nomura et al. 2004) has been widely investigated as a tumour marker in gastrointestinal malignancy and although it is commonly used in the diagnosis of pancreatic adenocarcinoma it has been evaluated in colorectal cancer. (Nakagoe, Sawai et al. 2003; Fuszek, Lakatos et al. 2004; Morita, Nomura et al. 2004; Chen, Yang et al. 2005) However, it again is not of use in diagnosis of early cancer although it may have a role in predicting poor outcome in colorectal cancer (Lindmark, Bergstrom et al. 1995).

Another novel protein found to be present in colorectal tumours and also in the serum of colorectal cancer patients at significantly higher levels than in controls, is Nicotinamide N-Methyltransferase, normally produced in the liver (Roessler, Rollinger et al. 2005). This has been shown to be more useful than CEA at detecting cancer from normal controls. The drawback of this marker is that it is produced in large quantities in liver disease and so may generate false positive results in this population. Further evaluation is needed for this to become acceptable as a clinical marker. Other tumour derived serum markers have been detected using proteomics to first identify proteins alpha-defensins 1 to 3 in solid tumour and an ELISA test then developed to detect the proteins of interest. This ELISA was then used on the serum of the patients who had the original tumour and gave a sensitivity of 68% and specificity of 100% for cancer (Melle, Ernst et al. 2005).

There are a variety of other serum markers, CA 72-4, β hCG (Louhimo, Alfthan et al. 2004), calcium (Fuszek, Lakatos et al. 2004) and ferritin (Griffiths and Schapira 1991; Scholefield, Robinson et al. 1998) as well as other more novel and unusual proteins that have been evaluated and current recommendations are that they are not of use in screening (Duffy, van Dalen et al. 2003). Where they may be useful, is in predicting response to therapy and perhaps in prognosis (Louhimo, Alfthan et al. 2004).

More recently, testing for circulating methylated DNA has attracted interest. As an example the Septin 9 test is now commercially available but it is associated with a sensitivity of 74% and a specificity of 90%(Warren, 2011).

1.1.6 Conclusions

Screening using FOB testing and colonoscopy is now established practice in many countries. There is no doubt that the FOBT is an imperfect test and although the reduction in mortality for screening with this method remains impressive, it is associated with appreciable false positive and interval cancer rates. Any screening test that achieves the twin demands of high sensitivity and high specificity as well as being cost effective and acceptable will be enthusiastically welcomed but as yet this remains elusive. The current goal of the screening programme is to reduce mortality from colorectal cancer and to eventually lead to fewer numbers of people presenting with advanced cancers. Population screening reduces the risk of colorectal cancer death occurring for a small number of individuals within large asymptomatic populations. Good uptake is crucial in order to achieve a reduction in deaths; the key to this is education for the public regarding colorectal cancer screening and, in contrast to other screening programmes which only identify malignant disease, there is the benefit of polypectomy in reducing cancer incidence.

Although the main aim of the screening programme is to reduce the mortality from colorectal cancer through early detection, any test that can reliably identify adenomatous polyps has the potential to reduce disease incidence. Efforts to find a better test have focussed largely on refining the FOBT, particularly by means of immunological technology and on the implementation of endoscopy based screening. However, increasingly sophisticated methods for detecting potential markers has opened up new fields and proteomics represents a promising avenue in this respect.

Chapter 1

Section 2

Clinical Applications of SELDI-TOF Proteomics

1.2.1 Novel Technologies

Novel proteomic technologies have the potential to advance our understanding of diseases, aid in the development of new strategies in therapeutics, permit scrutiny of organ, cell and intra-cellular protein expression and facilitate the search for novel biomarkers for the early detection of disease. Cancer is one of the leading causes of death in the developed world but although early diagnosis and intervention is associated with improved outcomes in cancer survival, simple diagnostic tests with high levels of sensitivity and specificity are lacking(Mayor 1998). Many tumour markers in current use are proteins which are either present or have their presence in the blood altered in the disease state; prostate specific antigen (PSA) is probably the most widely used but its limitations are well known(Carter and Isaacs 2004). From the earliest days of tumour markers, attention has been on the identification of informative proteins in easily accessible body fluids and the science of proteomics has opened up a new range of opportunities in this area(Rosenfeld 1987).

Proteomics can be defined as the large-scale analysis of the expressed protein complement of the genome(Feldman, Espina et al. 2004). The word proteome was first coined in 1994 and the human blood proteome is exceptionally complex, containing many different proteins, the most abundant being albumin. These proteins exist in various forms and sizes, existing in various states including precursors, mature forms, degradation products and differing degrees of post-translational modifications. Immunoglobulins, peptide and protein hormones, cytokines and other local mediators, tissue leakage products (creatinine kinase, cardiac troponins) as well as foreign proteins and aberrant secretions give rise to tens of

thousands of different proteins (Anderson 1995) (Anderson and Anderson 2002). With such a rich source of information easily available for sampling, interest has focused on searching for serum proteins that can offer simple, cheap, non-invasive tests for cancer and the coupling of bioinformatics with the new, high throughput mass spectrometry based tests has allowed rapid expansion in the field of biomarker discovery (Celis and Gromov 2003). The early detection of cancer, the monitoring of cancer treatment and individualization of therapy may all benefit from the discovery of as yet unknown biomarkers. Separation and identification of proteins is classically facilitated by two-dimensional polyacrylamide gel electrophoresis (2D PAGE) (Rabilloud 2002). The proteins of interest are made soluble and are then separated on the basis of charge over one dimension and molecular mass over the second. A current applied across the soluble proteins on a polyacrylamide gel strip allows the first stage of separation to occur and following further separation by mass, proteins can be visualized by staining with various techniques, such as coomassie blue and silver. Proteins of interest can be excised from the gels and subjected to sensitive and accurate mass spectrometric analysis and the mass spectral data is then used to interrogate specialized computer software protein identification databases (such as ProFound and Mascot) (Zhang and Chait 2000). These online data bases have allowed comparison of data and serve as a reference for investigators.

Mass spectrometers consist of an ion source, a mass analyzer that measures the mass to charge ratio and a detector to pick up the number of ions of each mass but peptides and proteins are difficult to measure by means of the usual mass spectrometry methods as they are often destroyed by the ionization process. (Aebersold and Mann 2003) 2D gel electrophoresis combined with mass spectrometry has limitations; in particular, it cannot accurately measure small differences in concentration and the traditional method of 2D

PAGE, despite some recent technical advances, is time consuming and not amenable to processing large numbers of samples(Rodland 2004). The results of tissue analysis by 2D-PAGE are greatly enhanced by laser capture microdissection (LCM) as this improves specificity and despite the disadvantages mentioned above, it remains a good means of detection of biomarkers(Banks, Dunn et al. 1999). Matrix-assisted laser/desorption ionisation time-of-flight mass spectrometry (MALDI-TOF MS) offers a more sensitive means of identifying small volumes of proteins and this rapidly became an accepted technology following its discovery. MALDI can determine the mass of a protein or peptide with a high degree of accuracy and it was quickly recognized that the masses of peptides generated from a pure fragmented protein with an enzyme of known size, such as trypsinogen, could uniquely identify a protein(Patterson 2003).

1.2.2.1 SELDI-TOF MS

Surface enhanced laser desorption/ionisation time of flight mass spectrometry (SELDI-TOF MS) offers a similar means of analyzing biological mixtures, the main difference between MALDI and SELDI being that in MALDI-TOF the surface or beads are passive probes and do not form part of the reaction whereas in SELDI-TOF, proteins from a biological sample are immobilized on one of a variety of chip surfaces, all with different binding specificities. A few microlitres of sample are bound onto the chemically treated surface of the chip by incubating the sample over a short time. The surfaces of the arrays have different specific properties; for example, the two ionic protein chip array surfaces consist of one strong cationic surface that binds proteins and peptides with a negative charge and the ionic surface which binds positively charged molecules. The arrays are 10mm wide by 80mm in length and have either 8 or 16 spot surfaces.



Figure 1.4

A CIPHERGEN ProteinChip showing the 8 spot array. Sample is bound onto the active chip surface

Washing the chip thoroughly after binding removes any impurities and unbound proteins. A matrix, such as α -cyano-4-hydroxycinnaminic acid (CHCA) which generates matrix ions less than 800 Daltons is then added which acts as an energy absorbing molecule. CHCA is particularly useful when peptides are being sought instead of heavier proteins as the matrix signal on the profile is at a low mass to charge ratio so that the signals from the peptides are not masked by the matrix noise. Sinnapinic acid, another matrix, generates higher mass to charge ratio ions, typically less than 2000 Daltons and is generally used in protein analysis. The matrix is applied to the chip surface in a saturated solution of 10% acetonitrile and 1% trifluoroacetic acid and as it dries onto the surface of the chip the proteins crystallize. The chip is then placed in a vacuum and a laser is directed on to the surface causing the desorbed proteins to be launched as charged ions by the addition of (usually one) proton. The vacuum chamber has an acceleration region and a free flight region, known as the drift tube and ions leaving the source are accelerated to the same energy through the drift tube by a voltage of 20kV applied across the surface of the chip. The time-of-flight (TOF) for each ion is analyzed to determine the mass to charge ratio (m/z), and as different masses have different velocities they arrive at the detector at different times. This is shown in Figure 1.5.

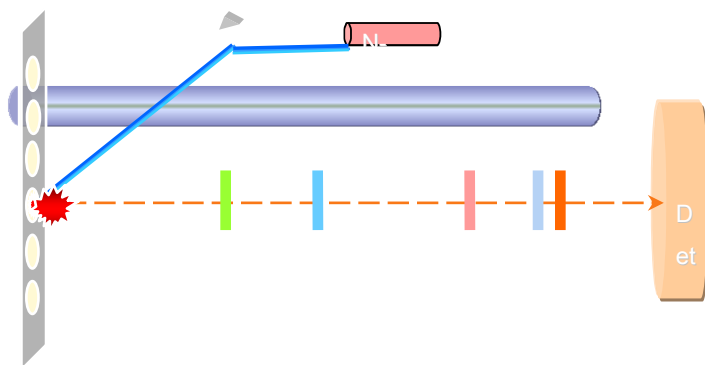


Figure 1.5

The chip is placed in the machine and a vacuum is created, a charge is applied across the surface and then a laser shone on the chip to lift the bound proteins. The particles then fly across the chamber and are detected at the other end of the tube where the time of flight is recorded.

The computer program attached to the analyzer then measures the time of flight and gives a mass to charge ratio. The equation shown shows how the mass charge ratio is calculated from the time of flight, as speed equals distance over time.

$$E = \frac{1}{2} mv^2$$

$$m/z = \frac{2V}{\Delta x^2} \quad \text{where } v = \frac{\Delta x}{\Delta t} = \frac{x_1 + x_2}{t_1 + t_2}$$

$$m/z = \frac{2V}{\Delta x^2} \cdot \Delta t^2$$

$$\text{Assuming } \frac{2V}{\Delta x^2} \text{ is constant}$$

$$V = \text{acceleration voltage}$$

$$v = \text{average final velocity}$$

$$m/z = K \cdot \Delta t^2$$

Equation for calculating time of flight

The accuracy of the machine is aided by regular calibration using samples of a known molecular weight and this can be done by adding known protein standards to the chip containing the sample (internal calibration) or running the standards through the machine on a separate chip (external calibration).

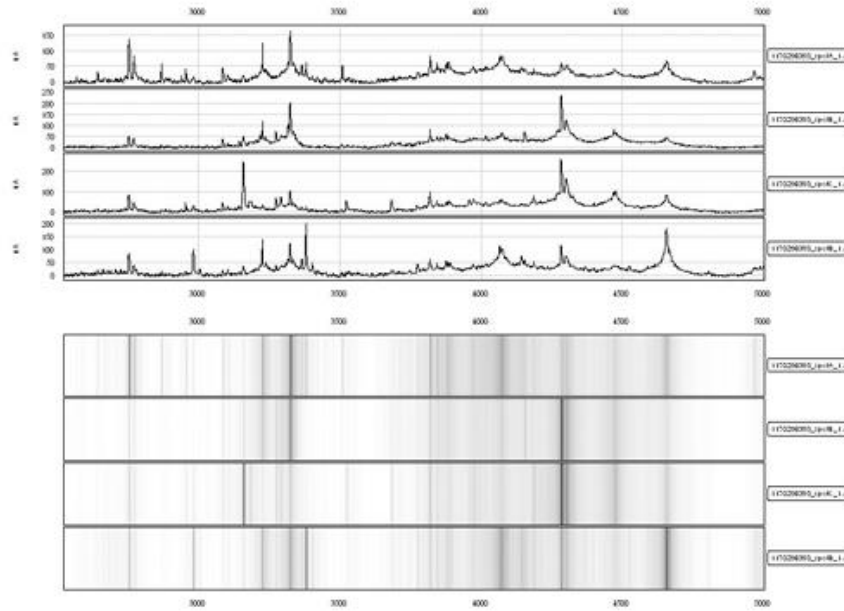


Figure 1.6

A typical SELDI-TOF profile showing peaks at mass/charge ratios which correspond to size in Daltons. The y axis shows the amount detected at the detector plate relative to other particles. Results can be viewed as a peak profile or as a gel view as shown in the figure above.

A typical SELDI-TOF proteomic profile will have anywhere up to 15 500 data points(Petricoin and Liotta 2004); these large amounts of data require dedicated proteomic bioinformatic data mining tools, which utilize two different methods of analysis. In the first system, the outcome is known ahead of time (so called “supervised” systems) and this type of system requires “training” by teaching the system which profiles represent diseased and which represent healthy states.

Proteomic patterns are analyzed by comparison to the training data and are given a score as to how much they resemble the original training data spectra. As the system looks at more and more spectra it can eventually learn to identify a new phenotype even if it has not been trained to see it beforehand. In the second “unsupervised” system, data are clustered into groups without previously being “taught” which are diseased. These systems are also

vigilant i.e. they learn and adapt as they analyze more data and there is a variety of pattern recognition tools that have been used for mining mass spectrometry data(Adam, Qu et al. 2002; Conrads, Zhou et al. 2003).

Sample fractionation, dividing complex biological mixtures into aliquots according to specific properties, can be used to improve the detection of low abundance proteins. SELDI accounts for the inherent variability of protein properties e.g. extent of hydrophobia, isoelectric point, by the different conditions of each surface and different spectra are generated from the same sample on different surfaces. The analysis of smaller abundance proteins, the subproteome, is aided by the removal of higher abundance proteins, e.g. albumin in serum samples. One of the main shortcomings of SELDI is that the relationship between the amount of substance present and the corresponding signal intensity is complex and not fully understood.

1.2.3.1 Proteomics in Diagnosis

Despite advances in understanding of the molecular basis of cancer the associated changes in protein profiles have been difficult to study until the advent of novel proteomic technologies. The dynamic nature of proteins, particularly in abnormal states, makes capturing changes particularly challenging but SELDI-TOF MS technology meets the challenges of both high sensitivity and high throughput. There has been much interest in proteomics as a platform for developing diagnostic tests, with particular reference to it as a screening tool in a wide variety of cancer types and the most promising recent studies have employed SELDI-TOF(Baak, Path et al. 2003; Rogers, Clarke et al. 2003; Gretzer, Chan et al. 2004; Wong, Cheung et al. 2004; Zhao, Gao et al. 2004).

1.2.3.2 Ovarian Cancer

In 2002 Petricoin et al reported a SELDI-TOF generated proteomic pattern that had a 100% sensitivity and 95% specificity for ovarian cancer serum detection(Petricoin, Ardekani et al. 2002). In this study, 50 cancer spectra were compared with 50 benign disease spectra and using an algorithm of the key low molecular weight marker proteins, a masked set of serum samples were used to identify cancer from non-cancer. The controls used were well matched and had other benign gynaecological disease or were “normal”. Understandably this paper generated a lot of interest and enthusiasm about SELDI-TOF MS (the authors won immediate praise from the press and the United States Congress passed a resolution to intensify this area of research(Congress 2002)) but it came under criticism from the mass spectrometrists and the scientific community (Sorace and Zhan 2003; Baggerly, Morris et al. 2004). The data made public by Petricoin’s group was reanalyzed and a series of problems centered on the baseline signal. This signal is highly variable from day to day and between machines and is usually due to matrix noise and to background electrical interference. Matrix noise is maximal below 2000 Daltons and the company that manufactures the technology (Ciphergen Biosystems, Fremont, CA) advocate ignoring protein peaks below this value. The work by Petricoin et al used peaks in this region to differentiate between the study groups and hence the interpretation of the results was based on unreliable data.

1.2.3.3 Head and Neck Cancer

In head and neck cancer SELDI-TOF proteomic analysis achieved 68% sensitivity and 73% specificity at detecting cancer from non-cancer in a study of 56 patients and 52 controls recruited from companions of the patients. Plasma profiles showed good reproducibility of the spectra on different days and on different chips opening up possibilities in this cancer

type. However, all the patients had quite advanced disease (stage III or IV) due to them being recruited from another study, so that the utility of this for screening as a screening test is not clear. (Soltys 2004).

1.2.3.4 Prostate Cancer

Several groups have looked at prostate cancer tissue which unsurprisingly, yields different spectra to normal healthy prostate samples (Zheng, Xu et al. 2003). Serum studies have been carried out aiming to detect biomarkers that could replace prostate specific antigen (PSA) and many different protein peaks have been identified using SELDI-TOF MS (Table 1(Hlavaty 2001; Adam, Qu et al. 2002; Petricoin, Ornstein et al. 2002; Qu, Adam et al. 2002; Ornstein, Rayford et al. 2004)). One of these groups went on to identify their biomarker in a subsequent publication, and when comparing cancer of the prostate to benign prostate disease, found their serum marker (which gave them 96% sensitivity) to be a vitamin D binding protein(Hlavaty, Partin et al. 2003). It is surprising that despite using the same chip surface, only two protein peaks were found in more than one study, these have been highlighted in bold. The groups conclude that the proteins that they identified are produced by the diseased prostate gland much the same way as PSA is leaked into the circulation by the prostatic columnar cells, but critics suggest that the molecules that are detected are related to the “epiphenomena” of cancer i.e. malnutrition, inflammatory cytokines, infection or cachexia (Diamandis 2003). What the differences make clear is that the methodology of sample handling and of analysis must be of extreme importance. The molecular mass of PSA is 27 755 Da but it was not detected by the SELDI-TOF MS analysis in these groups’ experiments (Belanger, van Halbeek et al. 1995). The proteins detected were in the lower molecular range and this calls into question the sensitivity of SELDI-TOF. Of particular

concern is that only those proteins which are abundant in the sample bind to the chip surface. PSA provides the perfect example of this as it is found at a concentration of between 4 and 20 μ g/L in the serum of prostate cancer patients whereas the overall total protein content is 3 to 5g/L.

1.2.3.5 Pancreatic Cancer

In one study patients undergoing pancreaticoduodenectomy for pancreatic adenocarcinoma had preoperative blood taken for SELDI analysis(Koopmann, Zhang et al. 2004). The control group consisted of patients with non-malignant disease of the pancreas (cysts, acute pancreatitis, chronic pancreatitis) and a subset of healthy people with no pancreatic disease. Four peaks were found that were significantly better than CA19-9, the currently used serum biomarker for pancreatic adenocarcinoma, at discriminating cancer from healthy controls and in the acute and chronic pancreatitis group they found different peaks that again were significantly better than CA 19-9. This study was particularly promising as the cancer group studied consisted of patients with small surgically resectable tumours who are potentially curable.

1.2.3.6 Breast Cancer

In breast cancer, mammography screening is well established but there are drawbacks to this modality as it delivers a dose of radiation and detects tumours only when they are large enough to cause an alteration in the mammogram image. Serum tumour markers for breast cancer (e.g. CA 15-3) are therefore being investigated although they lack the sensitivity and specificity required to be widely used in detecting early disease in populations(Ghnassia, Rodier et al. 2001). SELDI-TOF analysis of serum from 169 women, comparing cancer to benign breast disease has successfully discriminated cancer from non-cancer by using a panel

of three biomarkers(Li, Zhang et al. 2002). Two of these markers were up regulated and one was down regulated in patients with cancer and these were found to have a sensitivity of 93% and a specificity of 95% in discriminating between the cancer and normal or benign disease. No correlation was seen with amount of biomarker and tumour size or with lymph node status. Other, non-SELDI proteomic studies using MALDI and 2D-PAGE have found circulating serum proteins common to breast cancer patients which can discriminate between healthy and cancer patients serum with a 100% sensitivity and a 97% specificity(Rui, Jian-Guo et al. 2003).

1.2.3.7 Colorectal Cancer

Novel screening methods for colorectal cancer are required as the currently employed FOB tests lack sensitivity and specificity depending on the analytical sensitivity for haemoglobin and endoscopy is associated with poor uptake and is therefore of questionable value for population screening (Dachman and Yoshida 2003; Hardt, Toepler et al. 2003; Levin, Brooks et al. 2003; Whynes 2004).

A group in China searching for an identifiable colorectal cancer proteomic fingerprint in serum recently published their results(Zhao, Gao et al. 2004). Using the immobilized metal affinity capture chip (IMAC3) they analyzed serum from 73 cancer patients, 31 healthy patients and 16 cases of benign colorectal disease. They identified nine different markers and achieved impressive results with a sensitivity of 95.89% and specificity of 97.87%. The criticisms of this study are mainly that the control group included patients with adenomas and familial adenomatous polyposis. Another group, again in China, published results from a similar experiment that gave comparable results in colorectal cancer identification(Yu,

Chen et al. 2004); this group used a different chip surface to IMAC3(Zheng 2004) and a comparison of the two groups' results reveals no biomarker proteins in common, presumably related to different methodologies employing different chip surfaces. The key to accepting these biomarkers as diagnostic tests for colorectal cancer will rely on large numbers of patients and reproducible results corroborating the evidence from individual laboratories.

1.2.4.1 Other Screening Approaches

1.2.4.2 Cell Lines

Shiwa et al.(Shiwa, Nishimura et al. 2003) used cancer cell lysates from 39 different cell lines and analyzed them on five different chip surfaces. They identified a marker protein of 12kDa, prothymosin- α that was detected in high concentrations in colorectal cancer cells but not in normal colonic epithelium. Despite the small number of samples, four normal and four colorectal cancer, the cell extract profiles were striking and prothymosin- α was not identified in any other cancer cell line. Prothymosin- α is an acidic nuclear protein found in cells as opposed to serum so any implication as to its use as a screening biomarker is limited. However, in patients with metastatic adenocarcinoma of unknown primary there could be a role to differentiate one type of tumor from another in order to guide therapy. Villin, a microfilament associated protein that binds actin has been identified by proteomic profiling as being important in identifying colorectal from ovarian cancer in cell line studies.(Nishizuka, Chen et al. 2003)

1.2.4.3 Cervical cancer and Cytology

Cervical cytology smears have been explored by proteomic analysis and intracellular proteins from the exfoliated cervical cells have been used as a means of diagnosing invasive cancer from normal by using a discriminating pattern of seven proteins with a sensitivity of 87% and a specificity of 100%(Wong, Cheung et al. 2004). Both up regulated and down regulated protein peaks were found to be important but the proteins that were found to be useful were not identified.

1.2.4.4 Breast Cancer and Nipple Aspirate

Proteomic analysis of nipple aspirate fluid (NAF) in the diagnosis of breast cancer has been shown to yield a high accuracy for the detection of advanced disease (Sauter, Zhu et al. 2002; Coombes, Fritsche et al. 2003). The reproducibility of SELDI-TOF proteomics has been demonstrated by analyzing fluid from the same women over several days. A prospective trial that collected nipple aspirate fluid from 114 women undergoing diagnostic breast surgery aimed to determine protein masses associated with breast cancer whether or not there were any subsets of women with a unique proteomic profile and if this could be developed into a breast cancer predictive model. They concluded that the best cancer detection model included age, parity and the presence or absence of a 11,880 Da protein(Sauter, Shan et al. 2005).

1.2.4.5 Renal Cancers and Urine

Urine has been found to be of value in the diagnosis of urothelial malignancy(Celis, Wolf et al. 2000). Celis et al identified SELDI-TOF proteomic strategies that were used to identify metaplastic lesions in bladder squamous cell carcinomas, as well as biomarkers in the urine

for follow-up studies of patients with carcinoma. Renal cancer is often detected incidentally and is frequently advanced at the time of presentation with over half of patients having local or distant spread. Urine samples from patients before nephrectomy for renal clear cell carcinoma (RCC), from normal volunteers and from outpatients attending with benign diseases of the urogenital tract were successfully used to develop normal and diseased discriminating SELDI spectra (Rogers MA 2003). This was based on either the presence or absence of peaks or altered peak intensity values, resulting in sensitivity and specificity values of 98.3 and 100%. Work has also been done on serum in RCC patients, where the proteins have been identified as haptoglobin 1- α and serum amyloid α -1 and an as yet unidentified marker (Tolson, Bogumil et al. 2004).

1.2.4.6 Monitoring Treatment Response

In addition to the large body of work accumulating on diagnosis there is interest in using proteomic technology to examine response to therapy, particularly to chemotherapeutic drugs. In theory, treatment could be monitored by examining the proteomic profiles of serum both pre- and post-chemotherapy. The development of resistance to chemotherapy is thought to be caused by alterations of proteins within the cell. For example proteins associated with Vinca alkaloid resistance have been identified in human leukaemia cells by proteomic analysis (Verrills, Walsh et al. 2003) and in breast cancer patients undergoing chemotherapy with 5-fluorouracil, doxorubicin and cyclophosphamide, a single chemotherapy inducible serum protein peak was identified at 2790m/z when it appeared on day 3 in 80% of subjects undergoing chemotherapy (Pusztai, Gregory et al. 2004). Women receiving paclitaxel chemotherapy both adjuvantly and neoadjuvantly were studied and the protein was inducible in 80% of patients who were treated preoperatively and in approximately 40% of patients treated postoperatively but there was no clear correlation

between induction of the protein and final tumor response to preoperative chemotherapy. Five other peaks also were identified that discriminated between plasma from patients with breast carcinoma and plasma from normal women. These same peaks also were detectable in a subset of patients who already had undergone surgery to remove their tumors, suggesting that this could be used as a marker for residual disease.

In a study aimed at identifying proteases for drug targeting in colorectal cancer (McKerrow, Bhargava et al. 2000), tissue samples for 15 patients, one set of samples from the tumour, another set from adjacent normal colon and a last set from the liver metastases were analysed by SELDI-TOF. They discovered significant differences using SELDI-TOF in the levels of metalloproteases (high in the primary tumour, low in normal mucosa) and in mast cell proteases (high in normal mucosa, low in the tumour and absent in the metastases). The major proteases detected were metalloproteases; Cathepsin B activity was significantly higher in the primary tumour and even higher in the metastases.

1.2.4.7 An Update on SELDI-TOF MS

In the time since undertaking this research there has been a steady decline in the number of SELDI-TOF MS based publications. This is due in part to the evolution of more quantitative methods of Proteomics but is also due to the increased awareness of the drawbacks of this technique. The initial paper that employed SELDI to detect ovarian cancer with 100% sensitivity has yet to lead the development of a screening test for ovarian cancer and the results have not been reproduced by others.

The reproducibility of SELDI in different laboratories and also within the same laboratory has been studied, one group compared their own results in breast cancer with the published

literature and found that they had the same peaks as 25% of the published literature (Callesen, 2008). They later performed a reproducibility meta-analysis including their own peaks in ovarian cancer and had a slightly higher amount of same peaks (40%) that discriminated cancer from controls (Callesen, 2012)

In colorectal cancer there has been ongoing research in serum and also in tumour, adenoma and mucosal tissue using both SELDI and other proteomic techniques; Gelsolin is a down regulated protein in colorectal cancer tissue which was identified using iTRAQ proteomics (Fan, 2012).

In one very interesting paper serum samples were taken from the same patients (cancer and controls) on 8 separate days over 5 weeks and they were not able to identify cancer from controls when this variability was introduced (Albrethsen, 2012).

Wang et al had previously published results showing a 95% sensitivity and 95% specificity for a panel of peaks. Using the same serum samples that were shown to have previously detected peaks they used these samples in a more complex experiment, introducing a third group of general cancers and these peaks failed to discriminate CRC from other cancers but it was possible to separate normal from all cancers. (Wang, 2009)

Response to chemotherapy in colorectal cancer is attracting interest and predicting a good response to chemotherapy may be possible using a biomarker. Apolipoprotein A-I was shown to be associated with a good response to oxaliplatin and capecitabine chemotherapy (Helgason, 2010).

Table 1.7 summarises the recent work done in this field, the peaks are different in each study.

Author	Year	Tissue type	Aim of study	Peaks/Protein
Fan NJ	2012	Serum	Discriminate CRC from other cancer	741Da* 7772Da
Xu W	2014	Cancer, adenoma, mucosa	Discriminate cancer, adenoma, mucosa	Multiple peaks
Zhai XH	2012	Serum	Discriminate cancer from normal	3900Da
Fan NJ	2011	Serum	Node positive form node negative	3104Da, 3781Da, 5867Da, 7970Da, 9290Da
Lai Y	2008	Serum	Discriminate cancer from normal	8908Da, 13707Da

Table 1.7 Recent studies utilizing SELDI-TOF in colorectal cancer diagnosis.

*The peak at 741Da is likely to be due to matrix noise and should not be included in analysis.

In spite of the large amount of work done in SELDI-TOF based biomarker discovery in the 20 years since it was developed, there is not as yet a clinically applicable SELDI derived biomarker in use.

2.4 Conclusions

The field of proteomics is moving very rapidly but although many new technologies have been introduced over the past 10 years that have enhanced our knowledge in this field, developing a comprehensive library of the human proteome will be a considerably larger task than sequencing the human genome. Mass spectrometry based technology has established itself as an indispensable tool for interpreting complex samples but despite success in this field there are still significant technical challenges. There is much to commend SELDI-TOF

MS proteomic analysis; it is efficient at presenting information from a complex sample and interpreting patterns through its software analysis packages. However, the controversy generated by the ovarian cancer biomarker paper by Petricoin et al sounds a cautionary note(Petricoin, Ardekani et al. 2002). The reproducibility of SELDI-TOF MS is not established and conclusions drawn using proteins or fragments that have not been identified must be interpreted carefully. The other concern is that these proteins are not related to individual tumours but are part of the general inflammatory response: the so called epiphenomena of cancer(Diamandis 2003). There has been much emphasis on pattern recognition tools and matching in SELDI-TOF as it is this that holds the key to biomarker discovery (Petricoin, Paweletz et al. 2002) but there is still enough variability to treat any results with caution. What is clear from SELDI-TOF spectra is that a panel of up and down regulated proteins can discriminate between groups; this is a different concept to the single biomarker and even the term biomarker may need to be redefined to incorporate these new types of screening tools.

3.1 Aim of Study

- To collect serum and stool from patients with operable colorectal cancer and also obtain serum and stool from age and sex matched controls that we have proven do not have colorectal cancer (“clean” colonoscopy patients).
- To store in a database their characteristics and pathology.
- To develop a method of analysis of stool SELDI-TOF MS and apply this to the collected samples.
- To use SELDI-TOF MS technology to compare the proteomes of the two groups.
- To analyse colorectal adenocarcinoma and matched normal mucosa using SELDI technology and compare this to protein peaks detected in faeces and serum.
- To assess the future utility of these technologies as a screening tool

CHAPTER 2

MATERIALS AND METHODS

2.1 Clinical Methods

2.1.1 Patient Selection

Ethical approval was granted from the local ethics committee in accordance with the national guidelines on clinical research and sample collection commenced following this approval (December 2004). All the samples were obtained from patients after obtaining informed consent using the consent forms and information sheets contained in Appendix 2. There were two centres involved for collection of cancer samples, Ninewells Hospital and Medical School in Dundee and Perth Royal Infirmary in Perth, both part of Tayside University Hospitals NHS Trust. and all control samples were collected only from Ninewells Hospital. Control patients for this study were selected from patients who underwent a colonoscopy as part of the Scottish colorectal cancer screening programme pilot. Patients were excluded from the study if they had polyps, haemorrhoids and angiodysplasia. There are over 20 screening colonoscopies carried each month in Ninewells and from December 2004 to March 2006 patients were asked prior to colonoscopy to return a stool sample if they had a normal colonoscopy.

The cancer group was selected from those patients undergoing elective surgery for colorectal cancer; we excluded patients with known FAP, HNPCC or inflammatory bowel disease and all emergency presentations of colorectal cancer. Also excluded were all patients with known metastatic disease either pre-operatively or post-operatively. All diagnoses of colorectal cancer were confirmed post-operatively by pathological staging using the Dukes's classification and also the UICC (the International Union Against Cancer) Staging system.

2.1.2 Collection and storage of serum samples

All haemolysed samples were discarded from the study. Samples were all stored under the same conditions in the Tayside Tissue Bank, Ninewells Hospital. Serum was spun and separated by one of three specialist tissue bank staff within four hours of collection. Blood samples were collected in a red topped vacutainer (Serum Tube, Increased Silica Act Clot Activator, Silicone-Coated Interior) and were stored at room temperature until processing.

For the control group, conditions of collection were identical. The blood was collected by the same endoscopist each time and was drawn from a vein on the patients upper limb through a cannula inserted for the purpose of administering intravenous drugs just prior to colonoscopy. Colonoscopy requires cathartic bowel preparation and fasting prior to the investigation, 4 sachets of Kleen-Prep, (an oral powder, each sachet contains macrogol 3350 (polyethylene glycol 3350) 59g, anhydrous sodium sulphate 5.685g, sodium bicarbonate 1.685g, sodium chloride 1.465g, potassium chloride 743mg) must be taken prior to the test. Each sachet is reconstituted in 1 litre of water and approximately 250mL of reconstituted solution is taken every 10–15 minutes until all 4 litres have been consumed. This is commenced on the day prior to the investigation and the patient is fasted during this period. The solution from all 4 sachets should be drunk within 4–6 hours, flavouring such as clear fruit cordials may be added if required. Some patients were unable to tolerate such large volumes of fluid and for some it was necessary to replace the KleenPrep with Picolax, another oral powder which is sugar-free and contains sodium picosulfate 10mg/sachet, with magnesium citrate taken at a dose of 1 sachet in 200ml water in the morning (before 8a.m.) and a second in the afternoon (between 2 and 4p.m.) on the day preceding the procedure.

Cancer serum samples were collected from patients undergoing elective surgery for colorectal cancer. The serum samples were drawn when the patients attended for pre-operative assessment or on the immediate pre-operative day prior to any surgical intervention. The samples were collected by the admitting doctor on the ward or by NH and were drawn from a vein in the upper limb, or more rarely through a cannula. Blood samples were collected in a red topped vacutainer (Serum Tube, Increased Silica Act Clot Activator, Silicone-Coated Interior). Samples were then taken to the tissues bank for processing. 100 control samples and 86 cancer samples were collected.

2.1.3 Collection and Storage of Control Stool Samples

Stool samples were collected from patients following a clean colonoscopy result. In March 2006 patients who had undergone a normal colonoscopy in the preceding 16 months and had consented to be approached for a subsequent follow up stool sample were contacted by post and asked to return a stool sample. Samples were collected by the patients themselves in a blue universal container and the sample was then returned by first class post. 100 control stool samples were collected and 47 cancer samples.

The samples from cancer patients were collected pre-operatively, at the pre-operative assessment clinic the patients were asked to return with a stool sample when they attended the hospital for surgery. These samples were then collected by the patients and were then collected from the ward by NH and processed.

On receipt the samples were mixed with PBS and then the stool was left to stand for 30 minutes prior to vigorous manual homogenization. The stool saline mixture was then aliquoted for storage at -80C in the freezer until required. As all the control samples were recruited from a population that had tested FOB positive, the FOB status of each stool was

recorded at time of receipt. Most control stool samples were FOB positive, which led to them having the colonoscopy. These were used as the control samples as they all were proven not to have CRC and so in spite of the positive FOB were suitable for a control group.

Part 2 Laboratory based materials and methods

2.2.1. Solutions and Buffer Recipes

Buffers were made using reagents from Sigma-Aldrich unless otherwise stated. All de-ionised water was supplied by the department of surgery and molecular oncology. TRIzol was supplied by Invitrogen

General solutions

Phosphate buffered saline

Dissolve 8g of NaCl, 0.2g of KCl, 1.44g of Na₂HPO₄ in 800ml of distilled water. Adjust pH to 7.4 by titrating with HCl, make up to 1 litre with water.

Buffers for SELDI-TOF analysis

Buffers were made fresh with de-ionised H₂O from 1M stock solutions when required. This was calculated using the equation $C_1V_1 = C_2V_2$. Buffers were dissolved on a magnetic stirrer and pH was obtained and measure using a pH meter. The buffers were stored at room temperature and in daylight unless otherwise stated.

For the Q10 ProteinChip

Low stringency 50mM TRIS-HCl pH9

High stringency 100mM Na acetate pH6

To make a stock solution 1M TRIS-HCl dissolve 121.1g of TRIS base in 800ml of H₂O. Add concentrated HCl to the desired pH. Adjust the volume to 1 litre with H₂O.

To make a stock solution of 1M Sodium acetate dissolve 136.1g of sodium acetate into 800ml of H₂O. Adjust the pH with glacial acetic acid. Adjust the volume to 1 litre with H₂O.

For the CM10 ProteinChip

Low stringency buffer 100 mM Na Acetate pH 4.0

High stringency buffer 50 mM HEPES pH7

To make a stock solution of 1M Sodium acetate, dissolve 136.1g of sodium acetate into 800ml of H₂O. Adjust the pH with glacial acetic acid. Adjust the volume to 1 litre with H₂O.

To make a stock solution of 1M HEPES, dissolve 238.3g of HEPES into 800ml of H₂O. Adjust the pH to 7.0 with NaOH and adjust final volume to 1litre with distilled H₂O.

For the IMAC30 ProteinChip

Buffer: 100 mM Na phosphate pH 7.0

Pre-activation buffer 0.1M copper sulphate

Deactivation buffer 0.1M sodium acetate pH4

To make a stock solution of 1M sodium phosphate, dissolve 142.09 of sodium phosphate into 800ml of de-ionised H₂O. pH to 7.0 and make the solution up to 1 litre using de-ionised H₂O.

To make a 0.1M solution of copper sulphate, dissolve 15.9g of copper sulphate in 800ml of de-ionised H₂O. Add de-ionised H₂O to make up to 1 litre.

To make stock solution of 1M sodium acetate, dissolve 136.1g of sodium acetate into 800ml of H₂O. Adjust the pH with glacial acetic acid. Adjust the volume to 1 litre with H₂O.

For the H50 ProteinChip

Buffer: 0.1 % TFA

To make a 0.1% solution of trifluoroacetic acid, add 0.5ml of trifluoroacetic acid to 500ml of de-ionised H₂O.

Sample buffers

U9 Buffer (9M Urea, 2% 3-[(3-cholamidopropyl)dimethylammonio]-1 -propanesulfonic acid (CHAPS) 50 mM Tris-HCl pH 9)

In 80ml of de-ionised H₂O take 54.054g of Urea and add 2g of CHAPS (manufactured by BDH) and 6g of Tris-HCl. Carefully pH to pH9 with HCl and make up to 100ml with de-ionised H₂O.

2.2.1.2 Solutions and Buffers for tumour preparation

HEPES (1M pH7.5)

Dissolve 28.83g of HEPES in 80ml of de-ionised H₂O. Adjust pH with potassium hydroxide to 7.5 and make up to a final volume of 100ml with de-ionised H₂O. Store at 4C.

SDS PAGE and 2D PAGE Reagents

Sample buffer used was a solution of 4%SDS, 20% glycerol, 20mM Tris, 10mM EDTA, 200mM DTT with bromophenol blue at pH6.8.

The running buffer used was a solution of 1.92M with 250mM TRIS and 1%SDS.

Running gel recipe: 4ml H₂O, 3.3ml 30% acrylic mix, 1.5M TRIS (pH8.8), 0.1ml 10% SDS, 0.1ml 10% APS, 4uL TEMED.

Stacking gel recipe: 3.4ml H₂O, 830uL 30% acrylic mix, 630uL 1M TRIS (pH6.8), 50uL 10%SDS, 50uL 10% APS, 5uL TEMED

2.2.1.3 Experiments for optimisation of SELDI-TOF analysis

Preparation of stool samples for SELDI analysis

Stool samples were defrosted on ice and then mixed with U9 buffer. The stool and U9 mixture was vortexed for 30 minutes at 4C prior to profiling on the SELDI chip.

2.2.1.4 Preparation of the ProteinChip arrays

All chips were processed using these methods on an 8 spot array surface using the Cassette Compatible Bioprocessor (Cat. No. C503-0011).

Preparation of the Q10 ProteinChip Array (Strong Anion Exchanger)

Selection of appropriate binding buffers and sample buffers was performed. The ProteinChip array cassette was placed in the Bioprocessor and 200uL binding buffer was added to each well, the cassette was covered with parafilm and incubated for 5 minutes at room temperature (RT) with vigorous shaking (Micromix setting 20/7). The buffer was removed from the wells and immediately replaced with 90uL of binding buffer and 10uL sample. This was then covered and incubated for 60 minutes at RT with vigorous shaking on the Micromix. Samples were removed from the wells and washed with 200uL of binding buffer, with 5 minutes incubation with vigorous shaking on the Micromix again at RT; this was repeated for a total of three washes. The buffer was removed from the wells and 200uL of deionised water was added to each well and removed immediately, this was repeated once for a total of two water washes. The reservoir was then removed from the Bioprocessor and air-dried for 20 minutes until dry. Without removing the arrays from the cassette, 0.8uL of matrix was added and allowed to dry, a further 0.8uL of matrix was added and then allowed to dry.

Preparation of the CM10 ProteinChip Array (Weak Cation Exchanger)

Selection of appropriate binding buffers and sample buffers was performed. The ProteinChip array cassette was placed in the Bioprocessor and 200uL binding buffer was added to each well, the cassette was covered with parafilm and incubated for 5 minutes at room temperature (RT) with vigorous shaking (Micromix setting 20/7). The buffer was removed from the wells and immediately replaced with 90uL of binding buffer and 10uL sample. This was then covered and incubated for 60 minutes at RT with vigorous shaking on the Micromix. Samples were removed from the wells and washed with 200uL of binding buffer, with 5 minutes incubation with vigorous shaking on the Micromix again at RT; this was repeated for a total of three washes. The buffer was removed from the wells and 200uL of deionised water was added to each well and removed immediately, this was repeated once for a total of two water washes. The reservoir was then removed from the Bioprocessor and air-dried for 20 minutes until dry. Without removing the arrays from the cassette, 0.8uL of matrix was added and allowed to dry, a further 0.8uL of matrix was added and then allowed to dry.

Preparation of the H50 ProteinChip Array (Hydrophobic chip surface)

Selection of appropriate binding buffers and sample buffers was performed. ProteinChip arrays were placed in the cassette and then into a small bath containing 50% methanol for 5 minutes at RT. The methanol bath was then emptied and changed and the chips rewashed for a further 5 minutes at RT, giving a total of 2 methanol washes. Following this the cassette was placed in the Bioprocessor and 150uL binding buffer was added to each well, the cassette was covered with parafilm and incubated for 5 minutes at room temperature (RT) with vigorous shaking (Micromix setting 20/7). The buffer was removed from the wells

and repeated again, for a further five minutes incubation at RT on the Mocromix. The buffer was removed and immediately replaced with 90uL of binding buffer and 10uL sample. This was then covered and incubated for 60 minutes at RT with vigorous shaking on the Micromix. Samples are removed from the wells and washed with 200uL of binding buffer, with 5 minutes incubation with vigorous shaking on the Micromix again at RT; this was repeated for a total of three washes. The buffer was removed from the wells and 200uL of deionised water was added to each well and removed immediately. The reservoir was then removed from the Bioprocessor and air-dried for 20 minutes until dry. Without removing the arrays from the cassette, 0.8uL of matrix was added and allowed to dry, a further 0.8uL of matrix was added and allow to dry.

Preparation of the IMAC 30 ProteinChip Array (Immobilized Metal Affinity Capture)

Selection of appropriate binding buffers was performed. ProteinChip arrays were placed in the cassette and 50uL 0.1M Copper sulphate was added to each well and incubated for 10 minutes at room temperature (RT) with vigorous shaking (Micromix setting 20/7). The copper sulphate was removed from the wells and 200uL of de-ionised water was added to each well and incubated with shaking for 1 minute. The water was removed and a neutralisation buffer (200uL sodium acetate (pH4)) was added to each well and incubated for five minutes at RT on the Micromix. The buffer was removed and immediately replaced 200uL of de-ionised water which was again incubated for 1 minute at RT. The DI was removed and 200uL binding buffer was added to each well which was incubated for 5 minutes at RT. This was repeated once giving a total of two binding buffer incubations. This was then replaced with 100uL of sample (not in a sample buffer), covered and incubated for 60 minutes at RT with vigorous shaking on the Micromix. Samples are removed from the

wells and washed with 200uL of binding buffer, with 5 minutes incubation with vigorous shaking on the Micromix again at RT; this was repeated for a total of three washes. The buffer was removed from the wells and 200uL of deionised water was added to each well and removed immediately, this was repeated once for a total of two DI water washes. The reservoir was then removed from the Bioprocessor and air-dried for 20 minutes until dry. Without removing the arrays from the cassette, 0.8uL of matrix was added and allowed to dry, a further 0.8uL of matrix was added and allow to dry.

Preparation of the All-in-One Protein Standard II

All-in-One Protein Standard was prepared into 2uL aliquots which were kept in the freezer. When required a sample was removed from the freezer and defrosted, the sample was then spun in a microcentrifuge to collect the 2uL at the bottom and to this was added 8uL of EAM, this was then pipette mixed.

Preparation of the EAM matrix: Sinapinic acid

One tube of Sinapinic acid contains x ug of sinapinic acid, this was made into a saturated solution by adding to the tube 100uL of acetonitrile and 100uL of 1% trifluoroacetic acid. This was vigorously mixed on a vortex for 15 minutes and then briefly spun in a microcentrifuge to pellet any undissolved sinapinic acid.

2.2.1.1 Tumour sample analysis

Samples were supplied by the Tayside Tissue Bank and were taken from freshly removed surgical specimens that were frozen immediately following dissection. Samples are stored at -80 C in the tissue banks storage facility. Tumour samples were supplied as a 5mm cube of tissue with an equal sized piece of adjacent normal mucosa. Slides were taken from blocks of tumour and adjacent normal colon.

2.2.3 Laser Capture Microdissection

Tumour sections were taken from blocks of tumour and stained with haematoxylin and eosin and fixed onto slides 10 nanometres thick. They were then stained and dehydrated using the Arcturus HistoGene LCM Frozen Section Staining Kit. A total of 7 plastic jars were labelled and filled with

Jar a 75% ethanol

Jar b distilled water

Jar c distilled water

Jar d 7% ethanol

Jar e 95% ethanol

Jar f 100 ethanol

Jar g xylene

Using the solutions that were supplied with the kit, four slides at a time were processed through the staining and dehydration process. The sections were allowed to thaw for 30 seconds on a lint free towel. The slides were placed into 75% ethanol for 30 seconds and

then transferred to distilled water (jar b) for 30 seconds. The slides were removed and placed on the dry towel. 100uL of HistoGene staining solution was applied to the slide to cover the section and this was allowed to stain for 20 seconds. The slides were then placed into a distilled water wash for 30 seconds. The slides were then placed in sequence through increasing concentrations of ethanol for 30 seconds each and then finally into xylene for 5 minutes. The slides were dried and then transferred into a dessicator. Fresh solutions were used for each batch of four slides. They were then examined under the LCM System and a cap of 2000 cancer cells was collected along with matched normal mucosa. In accordance with the manufacturers instructions cells were added to 2UL of lysis buffer and the 8UL of PBS was added to make 10UL of lysate. The tube was inverted onto the cap and then centrifuged for 20 seconds at 13k. Samples were then ready for analysis. They were then examined under the LCM System and a cap of 2000 cancer cells was collected along with matched normal mucosa. In accordance with the manufacturers instructions cells were added to 2UL of lysis buffer and the 8UL of PBS was added to make 10UL of lysate. The tube was inverted onto the cap and then centrifuged for 20 seconds at 13k.

Protein extraction from tumour using SIGMA TRI reagent

Tissue samples were homogenised using the Ultra Turaxx Homogeniser in Trizol (1ml per 50-100mg tissue) and allowed to stand for 5 mins at RT to ensure complete dissociation of nucleoprotein complexes. Following this 0.2ml of chloroform per ml of Trizol used was added, covered and shaken vigorously and left to incubate for 15 minutes at RT. This was then centrifuged at 12000G for 15mins at 4 C. The resultant mixture was then in 3 phases, red organic (protein and DNA) interphase (DNA) and a colourless upper phase (RNA). Remove the aqueous phase and add 0.3ml 100% ethanol per 1ml of Trizol used. This was

was mixed and incubated for 3 minutes at RT and then centrifuged at 2000G for 5 minutes at 4C. The supernatant was removed for protein precipitation.

Proteins were precipitated from the supernatant with 1.5ml isopropanol per 1ml of Trizol and incubated for 10 minutes at RT, then centrifuged at 12000G for 10minutes at 4C.

The supernatant was discarded and the pellet was then washed 3 times in 95% ethanol (2ml per 1ml Trizol). Each wash consisted of a 20 minute incubation at RT for 20 and subsequent centrifuge at 7500G for 5mins at 4C to pellet the protein. 2ml 100% ethanol was then added and vortexed prior to incubation for 30mins at RT. Centrifuge for 5mins at 7500G at 4C and dry pellet under vacuum for 10 minutes. The pellet was then dissolved in Triton X100 with repeated pipetting and given a final centrifuge at 10000G for 10mins at 4C and used immediately.

2.2.4 Depletion of serum

Serum was depleted of albumin, transferrin, Haptoglobin, Anti-trypsin, IgG and IgA, using the Agilent high capacity multiple affinity removal spin cartridge. Serum samples were defrosted on ice and 14uL was diluted with buffer (named as “Buffer A” in the kit.) to a volume of 200uL. This was filtered through a 0.22um spin filter and collected. The spin column was then prepared by washing through 4ml of Buffer A, excess buffer was removed from the top of the resin bed to allow for the filtered, diluted serum to be added. The sample was then added and the spin column capped loosely, a collection tube, labelled flow through fraction 1 (F1) was attached to the bottom of the collection tube and this was centrifuged gently for 1 minute 30 seconds. The cap was removed and 400uL of Buffer A was added to the top of the spin column, this was then centrifuged for 2 minutes 30 seconds at 100 x g and the flow trough was collected in the F1 tube. The F1 tube was removed, the collected flow through was stored and replaced with a fresh collection tube (F2), a further 400uL of Buffer A was added to the top of the column and this was then centrifuged for 2 minutes 30 seconds at 100 x g. The collected flow through was stored. The bound proteins were then eluted from the spin column resin by slowly pushing 2ml of Buffer B through the cartridge and into a fresh collection tube (Bound) and following this elution the spin column was re-equilibrated by slowly pushing 4ml of Buffer A through the spin column, this was not collected. The spin column was then re-used or returned to the fridge (4 C) for storage.

Both the sample tube and the sample were chilled for 30 minutes at 4C. The sample was then mixed with acetone at a ratio of 1:6 and the tube was inverted 3 times. Following this the sample was chilled at -20C and observed carefully for formation of a flocculent which occurred between 1 to 4 hours. The sample was then spun at 6000 x g for 10 minutes and the acetone was decanted and not allowed to dry.

2.2.5 Miscellaneous experiments

The Bradford Assay

This is the standard and traditional method of assessing protein concentration of a sample.

Micro Assay Procedure

The spectrometer was switched on and allowed to warm up for 15 minutes before use. Standards were prepared using Bovine Serum Albumin over a range of concentration from 1 to 20 micrograms of protein to a volume of 200 μ l in each cuvette. Unknowns were also prepared to estimated amounts of 1 to 20 micrograms protein per tube to a final volume of 200 μ l. 800 μ l dye reagent was added to each and incubated for 5 minutes before the absorbance at 595 nm was measured and noted down for each standard and each unknown.

Macro Assay Procedure

The spectrometer was switched on and allowed to warm up for 15 minutes before use. Standards were prepared using Bovine Serum Albumin over a range of concentration from 20 to 200 micrograms of protein to a volume of 200 μ l in each cuvette. Unknowns were also prepared to estimated amounts of 20 to 200 micrograms protein per tube to a final volume of 200 μ l. 800 μ l dye reagent was added to each and incubated for 5 minutes before the absorbance at 595 nm was measured and noted down for each standard and each unknown. When large numbers of samples were being processed simultaneously, the assay was performed in a microwell plate and the sample made up to a total volume of 200 μ l in each well.

Protein analysis by SDS PAGE

SDS-Page sample buffer was added to the protein sample and then heated to 95°C for 5 minutes therefore denaturing the proteins. These samples were then loaded onto a 10% polyacrylamide gel and resolved by electrophoresis in running buffer in an apparatus at 150V for 45 minutes. A control lane containing 2.5µL of prestained, coloured low molecular weight markers (check maker) was included. The gels were then fixed in fixing solution for 10 minutes and then a coomassie stain was added for a further 30 minutes with incubation on a rocker at room temperature. When the gels had fully stained, they were de-stained on the rocker with several changes of de-staining solution and a small piece of crumpled up paper towel to absorb the released coomassie stain. The gels were then dried using the gel drying apparatus and photographed using a digital camera.

FOB Testing

On receipt of the stool sample, an FOB test was performed using one Hemocult II testing kit per sample. As per the manufacturers instructions, a small piece of stool was smeared onto both windows covering them entirely and then immediately developed using the Hema-Chek developer (stabilised hydrogen peroxide in aqueous methanol, Bayer Corporation, IN, USA). The window was then assessed for a colour change from white to blue, indicating a positive result. Any amount of blue was assumed to be positive.

CHAPTER 3

ANALYSIS OF COLORECTAL CARCINOMA AND NORMAL MUCOSA

3.1 Introduction

Colonic epithelium renews itself every five to six days, and suffers extreme environmental stress from intra-luminal contents. Colorectal tumours are acquired through genetic mutations that confer a growth advantage to the cell characterised by a change in epithelial morphology; from normal into dysplasia, followed by adenoma and through to carcinoma(Conlin, Smith et al. 2005). In this mucosa to adenoma to carcinoma sequence there is a mutation of several genes which trigger cell proliferation, growth and transformation and prevention of cell cycle arrest. These changes are multiple and varied and colorectal cancer is molecularly and genetically heterogeneous(Hardy, Meltzer et al. 2000). Disturbances in the intracellular proteins may be then translated into changes in serum or stool. For this reason, tumours were analysed with SELDI-TOF MS to identify if there were any peaks present in the solid tumour that would also be seen in the peak profiles found in the serum and stool studies.

3.2 Optimisation of method

Initially, four pieces of colonic adenocarcinoma and matched normal mucosa from the Tayside Tissue Bank were used as the optimisation tumours. Samples were prepared by the tumour bank who removed samples of tumour and normal mucosa from fresh specimens and these were then stored at minus 80 C.

A comparison of SELDI profiles obtained by tissue homogenization

The tumours for homogenization weighed between 0.69g and 1.24g with an average weight of 1.01g (when frozen at -80C), the matched normal mucosa weighed between 0.98g and 1.82g with an average weight of 1.60g. The optimisation samples (tumour tissue matched with normal mucosa) were then prepared according to the protocols.

Two different buffers were used for sample preparation for the solid tumour and these were then compared on different chip surfaces (CM10 at pH4 and pH7 and IMAC30). The CM10 chip profiles at both the low and neutral pHs were disappointing and this surface was discounted as a candidate for profiling (Figure 3.1).

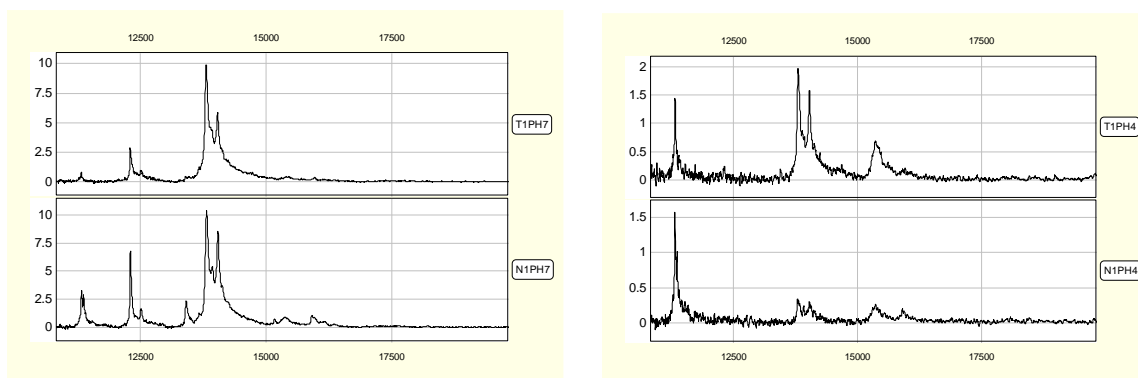


Figure 3.1 Sample incubated onto the CM10 chip surface at pH7 and at pH4 did not show useful SELDI profiles. The peaks are broad and of very low abundance.

The IMAC 30 chip revealed clean spectra and this was investigated further at varying concentrations of sample; 10 microlitres, 20 microlitres, 50 microlitres and 70 microlitres were incubated onto the chip surface in the buffer. The results of this are shown in Figure 3.2. An optimal amount of 50 microlitres of sample was decided on and this was used in all the experiments, there was no appreciable difference in the mean peak intensities of each protein peak at the differing concentrations.

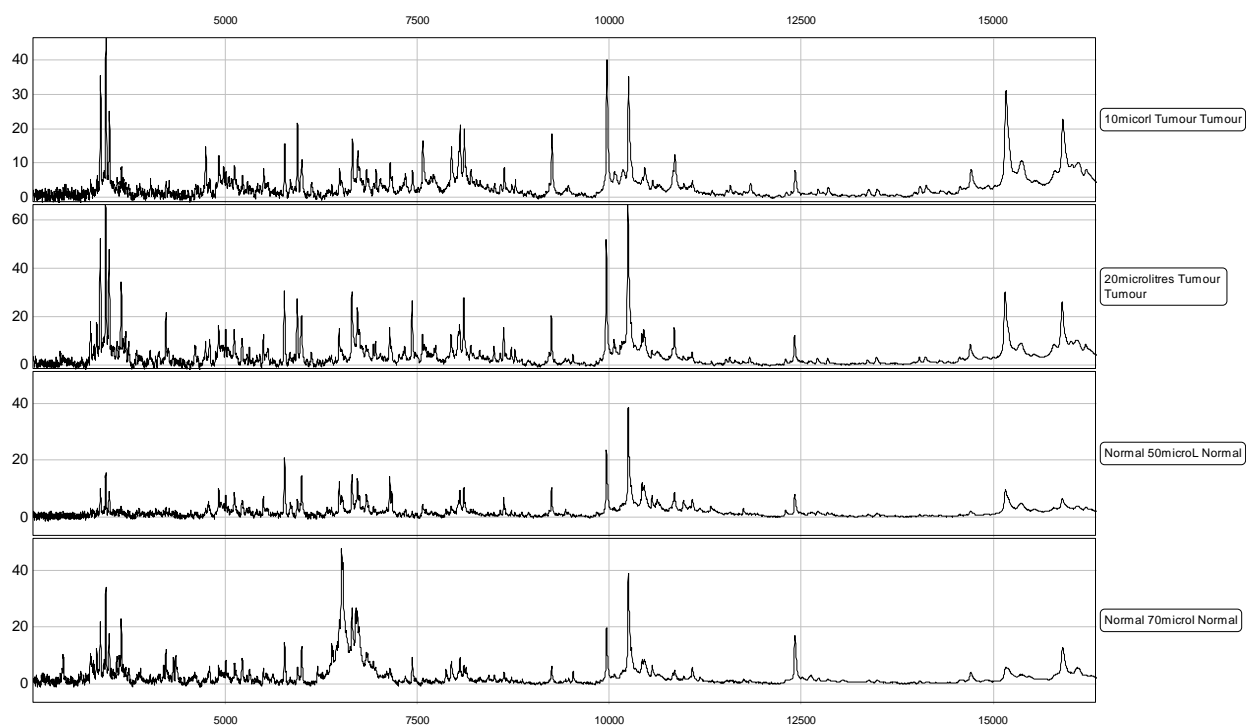


Figure 3.2 This figure shows the profiles obtained from two samples of tumour at differing concentrations in the binding buffer and two samples of normal tissue at two different concentrations. Varying the amount of sample in the buffer did not dramatically impact on the mean peak intensities obtained or the number of peaks. The normal sample at 70 microlitres has an increase in the binding of the peak at 6500Da and 12500Da but the peaks are present in both the higher and lower concentrations at detectable levels.

3.3 Analysis of tumours with matched normal mucosa using the IMAC30 chip

Following selection of the optimal chip surface, eight tumours from tissue bank were compared to matched normal mucosa. All were operative resected specimens and so were presumed to be potentially curative at the time of surgery. The details of the tumours analysed are shown in Table 3.1

PATIENT	SEX	GRADE OF DIFFERENTIATION
001	f	Well
002	f	Moderate
003	f	Moderate
004	m	Moderate
005	f	Moderate
006	f	Moderate
007	m	Moderate
008	m	Moderate

Table 3.1 Patient sex and degree of differentiation of the tumours analysed on IMAC chip with matched normal mucosa.

The Biomarker Wizard software data analysis package was employed for data analysis. The data generated from each chip was imported into one experimental file and the baseline was subtracted from each spectrum and all data were normalized with a total ion current between 1500 to 20000Da using a calculated normalization factor of 1.99. All peaks were then detected using the automatic peak detection facility, a ratio of 10 signal to noise was selected with a minimum peak threshold of 30% of all spectra. In total 54 peaks were identified as being significantly different between the two groups (126 peaks were detected in total). Four of these were less than 2kDa and so were assumed to be due to matrix noise. The mass to charge ratios were rounded up to a whole number. The mean peak intensities were then rounded up or down one decimal place as were the standard deviations. There were 19 M/Z

ratio peaks which were higher in the tumour group and 31 peaks which were higher in the normal group (Table 3.2). Comparison of tumour and matched normal mucosa revealed differences in both the up and down regulation of protein peaks.

MASS TO CHARGE RATIO (DA)	P VALUE	MPI NORMAL	MPI TUMOUR
2023	0.001	4.2	0.6
2055	0.001	3.2	0.9
2186	0.001	7.7	2.6
2205	0.001	2.2	0.1
2352	0.001	2.9	0.0
2378	0.012	1.4	0.3
2486	0.015	4.4	2.6
2582	0.004	2.2	0.4
2625	0.015	1.4	1.3
2735	0.012	2.2	1.4
2826	0.001	0.0	3.5
2921	0.031	1.4	2.5
2971	0.004	3.9	0.9
2999	0.015	2.8	1.8
3031	0.001	2.9	0.4
3093	0.001	18.9	4.7
3250	0.031	1.5	7.7
3296	0.012	6.6	4.0
3359	0.004	5.0	4.3
3374	0.001	4.1	17.7
3410	0.001	3.3	6.2
3444	0.001	3.5	21.1
3489	0.001	4.5	15.4
3543	0.001	5.1	2.9
3563	0.001	4.0	2.0
3853	0.012	3.7	1.2
3913	0.001	4.3	1.0
3944	0.004	2.8	0.5
4054	0.012	3.5	1.5
4094	0.012	0.7	2.2
4126	0.001	1.7	4.1
4228	0.031	1.4	8.9
4493	0.012	2.6	1.7
4753	0.001	2.8	0.8
4922	0.004	2.8	1.0
5162	0.001	3.9	0.7
5261	0.004	5.5	0.9
5289	0.001	4.9	0.8
5497	0.001	0.9	4.0
5686	0.012	1.7	4.2
5717	0.002	1.5	3.9
6054	0.001	2.9	0.6
7565	0.015	10.2	7.6
7748	0.001	6.3	2.6
7793	0.001	6.3	1.4
8056	0.001	9.8	2.5
8793	0.001	2.7	0.4
9144	0.001	1.8	0.1
10853	0.001	0.3	4.4
16095	0.001	7.5	8.7

Table 3.2 Peaks identified by biomarker wizard as different between the two groups. This table shows the mean peak intensities (MPI) for each type and the P values. The peaks that are upregulated in cancer are shown in bold.

The profiles were then individually assessed for the presence of double and triple charged peaks. There were three peaks which differed markedly between cancer and non cancer. Average peak intensity for the 3372Da biomarker was 17.7 in the tumour sample compared to 4.1 in the normal mucosa. The average peak intensity for the 3443Da biomarker was 21.1 in the tumour sample compared to 3.5 in the normal mucosa. Figures 3.3, Figure 3.4, Figure 3.5 and Figure 3.6 show the SELDI spectra for 8 pairs of cancer and matched normal mucosa over different mass ranges. The biomarker peaks are identified in red in the cancer samples. The peaks were not present in all eight of the tumours. The mean peak intensities were inspected. Upregulated peaks in the tumour group were of most interest, as these may be found in the serum and stool.

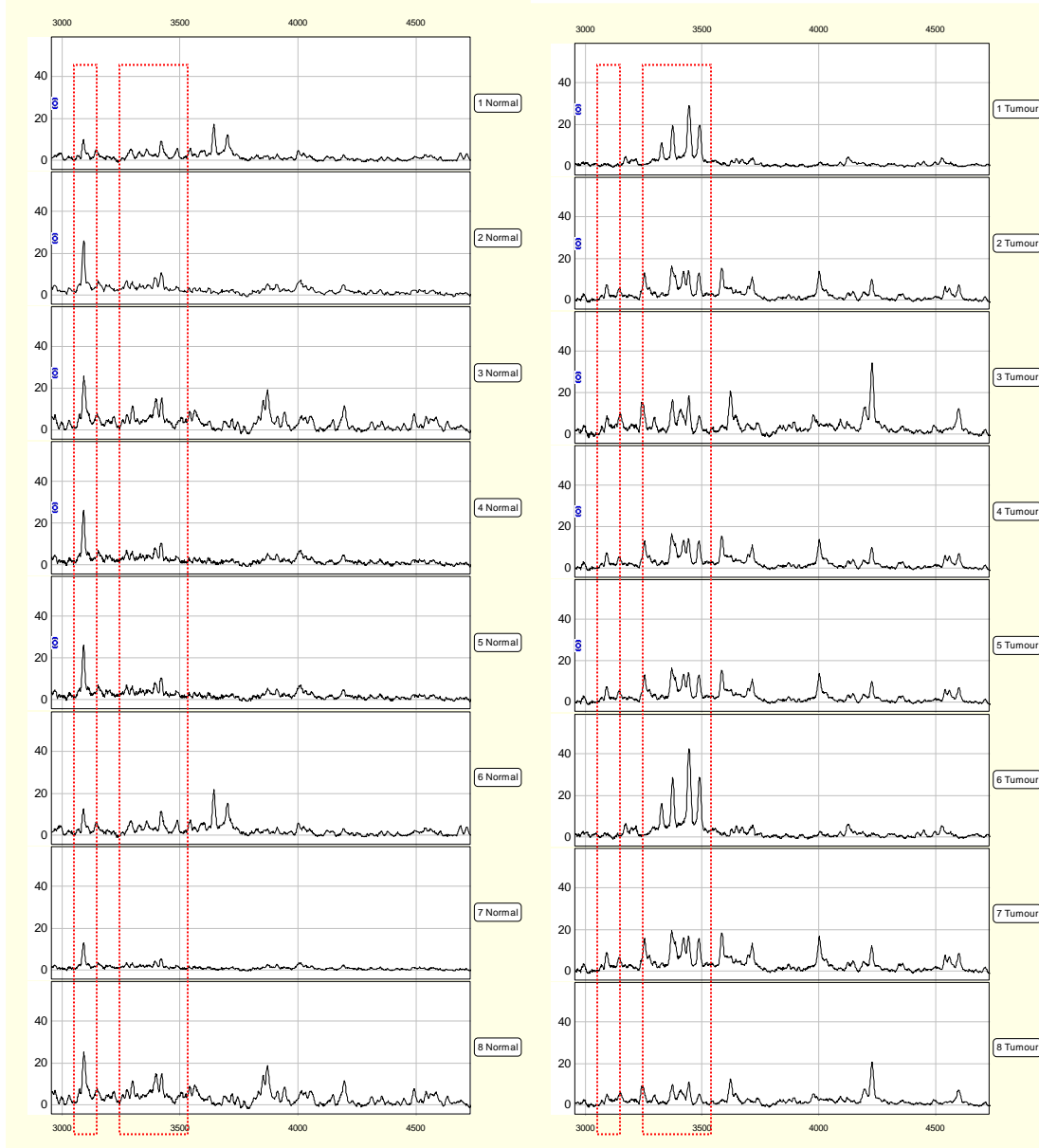


Figure 3.2

This shows the peaks detected in the 3kDa to 4.5kDa mass range. The up regulated peaks of interest in the tumour samples, 3374, 3444 and 3490 Da are circled in the red dotted square. In the normal sample the same three peaks are not present. In the mucosa samples, the upregulated peak that is seen at 3093Da is not seen in the tumour samples.

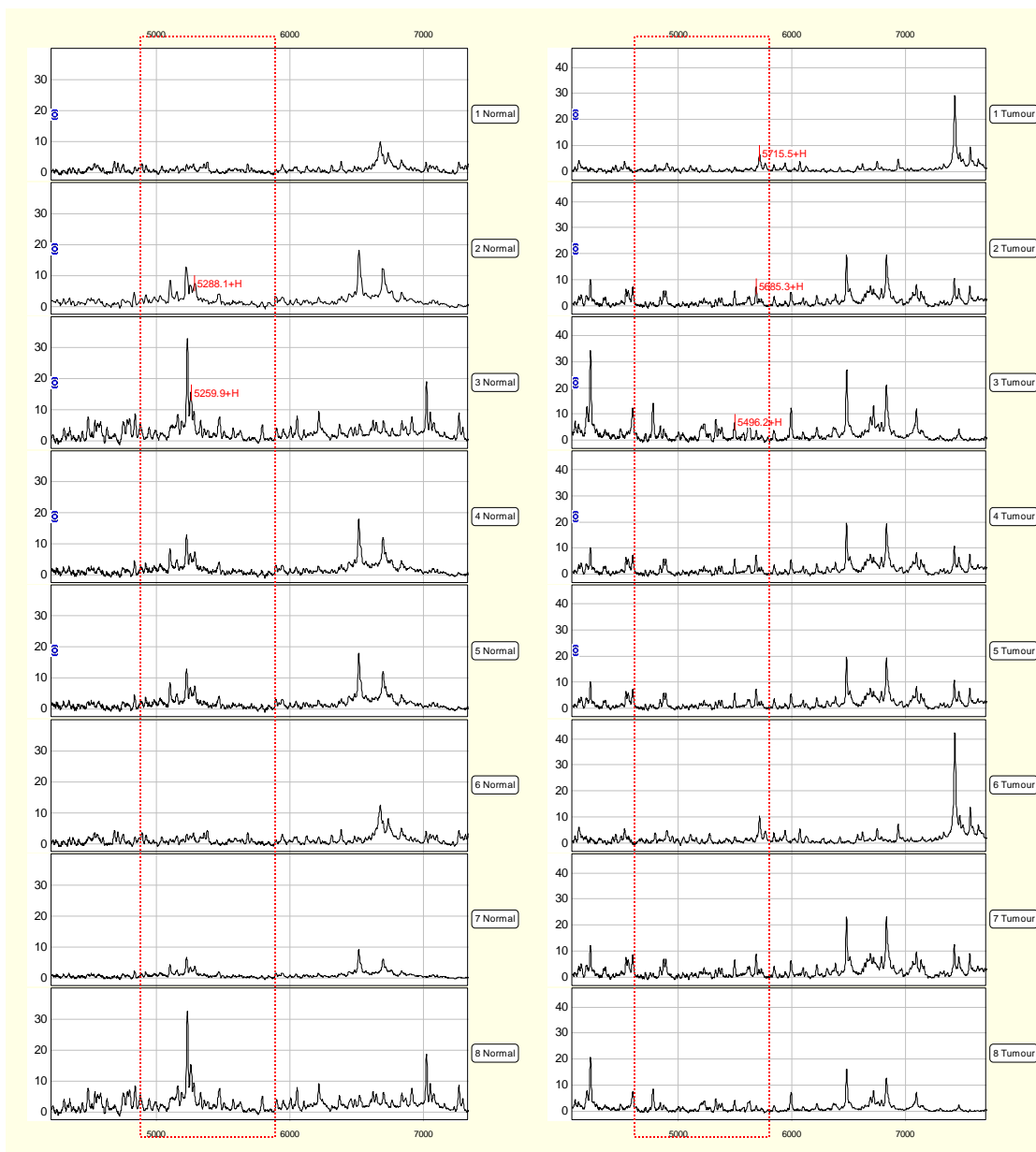


Figure 3.3

This shows the peaks detected in the 5kDa to 7kDa mass range. The peaks of interest are labelled, in the tumour samples, 5497, 5686, 5717 Da are present and in the mucosa samples, the peaks that are upregulated are at 5261 and 5289Da. The area of interest is highlighted in the red box. Note that the Y axis scale is labelled differently in the tumour compared to normal (max of 30 compared to 40).

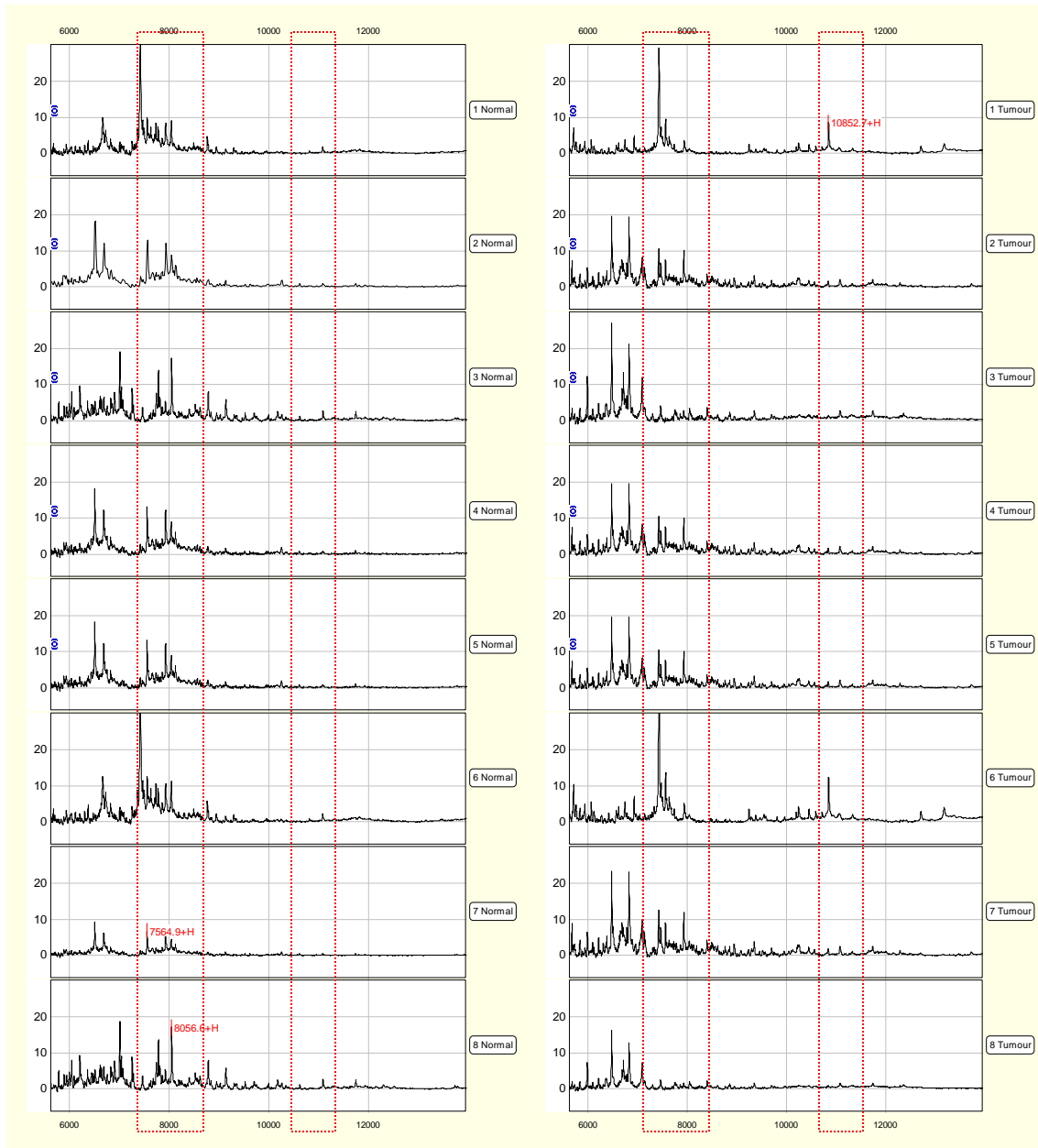


Figure 3.4

This shows the peaks detected in the 6kDa to 12kDa mass range. The peaks of interest are labelled. In the tumour samples, there is a peak at 10854Da and in the mucosa samples, the peaks that are of interest are at 7565 and 8056 Da. The area of interest is highlighted in the red box.

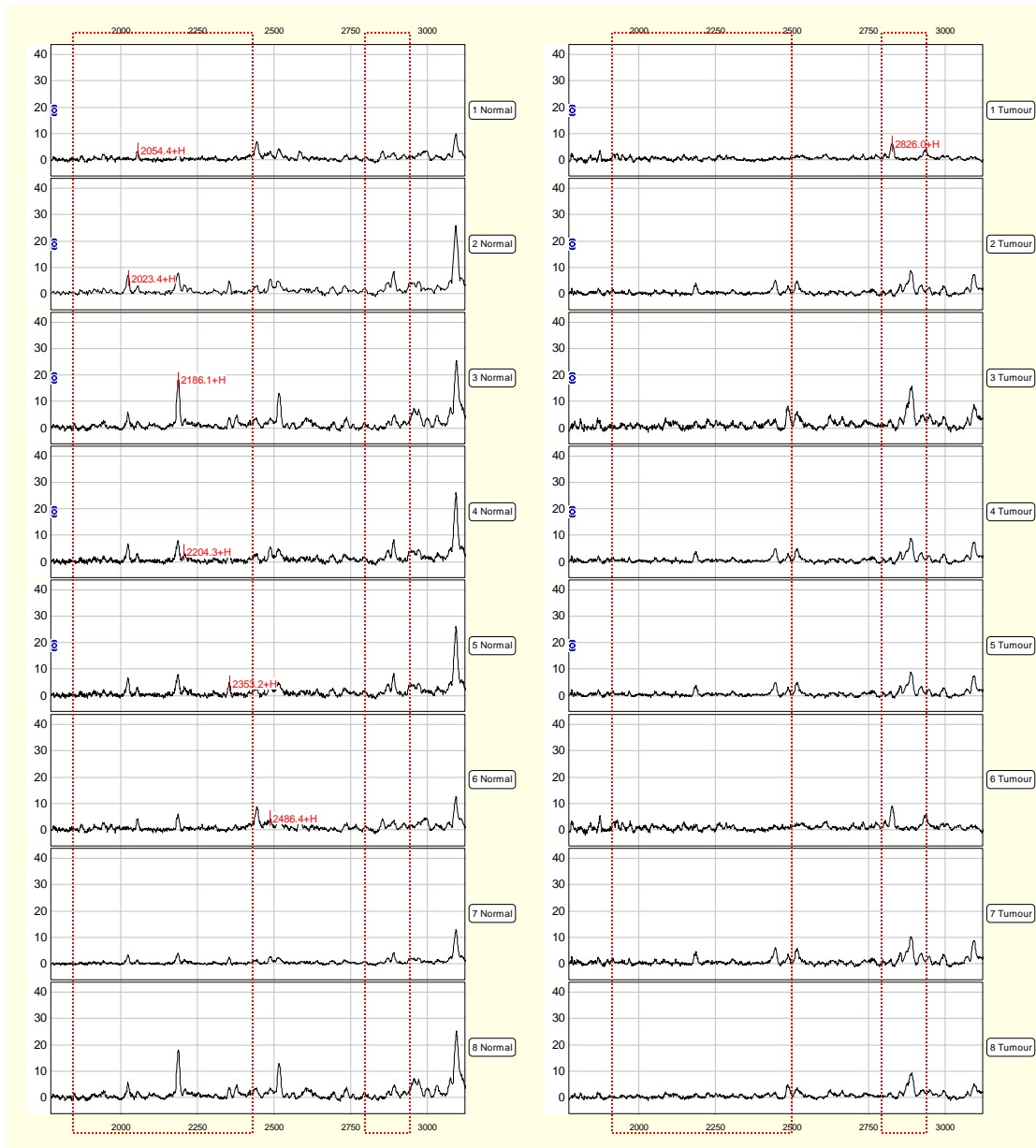


Figure 3.5

This shows the peaks detected in the 2kDa to 3kDa mass range. The peaks of interest are labelled. In the tumour samples, there is a peak at 2826Da that is not present (MPI of -0.4) in the normal samples. In the mucosa samples, the peaks that are of interest are at 2023, 2055, 2186, 2205, 2352, 2378, 2486, 2582, Da. The area of interest is highlighted in the red box.

3.4 Laser Capture Microdissection combined with SELDI-TOF analysis

The frozen sections (Pathology Department) for LCM were prepared and dried according to the protocol, caps were used to select 1000 cells from both the cancers and normal and these were lysed into a buffer. Preparation of the samples on the CM10 chip and the IMAC 30 chip revealed a disappointing trace and the protein concentration for each sample was checked and found to be below the optimal threshold for SELDI profiling. It was estimated that 2000 cells would be required for this method to yield sufficient proteins for analysis.

CHAPTER 4 ANALYSIS OF HUMAN SERUM

4.1 Introduction

Screening for colorectal cancer currently involves a series of gFOB tests and an invasive, expensive colonoscopy that carries a small risk of significant harm from complications. If a serum biomarker could be identified that would allow identification of patients with colorectal cancer, or even more importantly, those with high risk adenomas then this process would be revolutionised to a simple blood test that could reduce the numbers of negative colonoscopies considerably. The development of such a test would necessitate the identification of a biomarker and subsequent development of an antibody.

There is a move towards personalised medicine and genomics and proteomics are at the forefront of this. Detailed study of an individual's tumour, both genetically and in the systemic response is leading to increasingly personalised treatment. How applicable this type of information is in the field of screening for cancer where cost, acceptable levels of sensitivity and specificity and a proven improvement in public health are paramount remains to be seen. Serum proteomics using SELDI-TOF MS requires very small amounts of sample at low protein concentrations, in contrast to the traditional methods that require large amounts of protein to obtain meaningful results and throughput of samples is relatively high. Serum is rich in protein, in the normal adult concentration is 40-70 mg/mL protein, but it is dominated by a small number of very abundant proteins; up to 70% of serum is albumin, the other major components are transferrin, IgG, haptoglobin, alpha-1-antitrypsin and Ig A. The other proteins, are numerous in number but very small in concentration, spanning perhaps 10 orders of magnitude of concentration. Albumin is a sticky protein and many smaller protein fragments and proteins are bound to it.

Pie chart representing the constituent proteins in serum

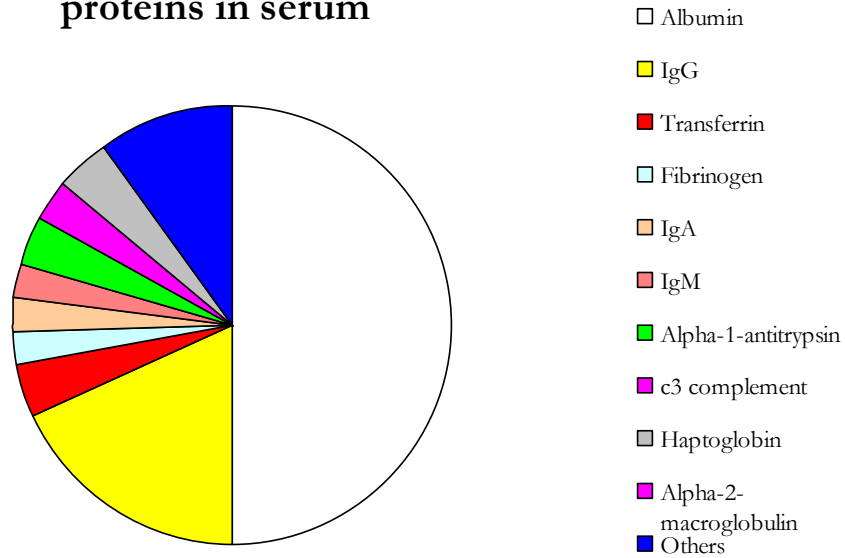


Figure 4.1 Pie chart showing constituents of human serum

Within the “others” group lies all the circulating proteins, enzymes, transport and binding proteins, complement and clotting factors (more abundant in plasma than serum) transport and binding proteins, proteases, protease inhibitors, cytokines, growth factors, hormones and an exhaustive list of intracellular and extracellular proteins, existing in a huge variety of precursor and degradation states. This is shown in Figure 4.1.

In screening for colorectal cancer, it would be ideal to detect early cancers and adenomas when the chance of any treatment being curative is the highest. However, when disease is at an early stage the systemic effects of the cancer and thus the detectable changes in the serum are likely to be low. Any changes that are detectable are likely to be small and multiple and SELDI-TOF analysis is well placed to identify these.

4.2 Results

Samples were collected from the two groups, cancer (n=100) and normal (n=100).

There were more males than females in the cancer group (71% male versus 29% female).

The age distribution of the cancer group was older than the screening group, average age of 75.4 years (48 to 79 years) compared to 66.1 years. The majority of patients in the cancer group had presented symptomatically and were not screen detected, a number of patients approached were excluded from the study post-operatively as they were found to have metastatic disease of the liver (n=5) or to have had a benign polyp (n=3). None of the serum patients were from the tumour group of samples.

Pathological information on tumours

The patients recruited into this study included rectal cancers, sigmoid colon cancers, descending colon cancers and right sided cancers. There were more rectal cancers than any other site (54%). Sigmoid and descending colon cancer accounted for almost a third of the group and the remaining cancers were on the right side of the colon (caecum and hepatic flexure). The degree of differentiation varied from 76% moderately differentiated and the remaining cancers being either poorly differentiated (19%) and well differentiated (5%).

There were no poorly differentiated cancers less than a T3. 50% of right sided cancers were poorly differentiated compared to just 18% of rectal cancers and none of the sigmoid cancers were poorly differentiated. There were no N0 patients that were poorly differentiated compared to 50% of the tumours that were staged as N2.

Staging of cancers was done by two methods, Dukes and the UICC TNM method. The majority of cancers were Dukes C, 26% were C1 and 24% were C2. Only 16% of cancers were Dukes A and 34% were Dukes B.

4.3 Serum analysis

Initially a test set of samples was used to optimise the method of analysis, these were the first ten consecutive samples collected and were a mix of cancer stages. The controls were also randomly selected from the initial control samples collected from within the screening programme (age range 50 to 74 years). The protein concentration of 10 normal serum samples was measured using the Bradford method and found to vary from 48mg/ml to 80mg/ml with a mean value of 66mg/ml. The protein concentration of 10 cancer serum samples was measured using the Bradford method and found to vary from 36mg/ml to 76mg/ml with a mean value of 68mg/ml. Each group contained five males and five females. The control group were all fasting and it is not known how many of the cancer group were fasting.

SDS PAGE analysis of serum

A serum sample was then examined using the traditional method of SDS PAGE and stained using coomassie blue. The results are shown in the figure 4.2. There was a large smear from the abundant proteins, in particular a protein at 66kDa which was highly abundant and likely to be albumin. The limitations of discovering biomarkers using this type of technology can easily be appreciated. Even after depletion of the sample for Albumin, IgG, IgA, transferrin, alpha-1-antitrypsin, haptoglobin and albumin the gel was difficult to read and small abundance proteins were not visible using this technique as the amounts of protein is too small.

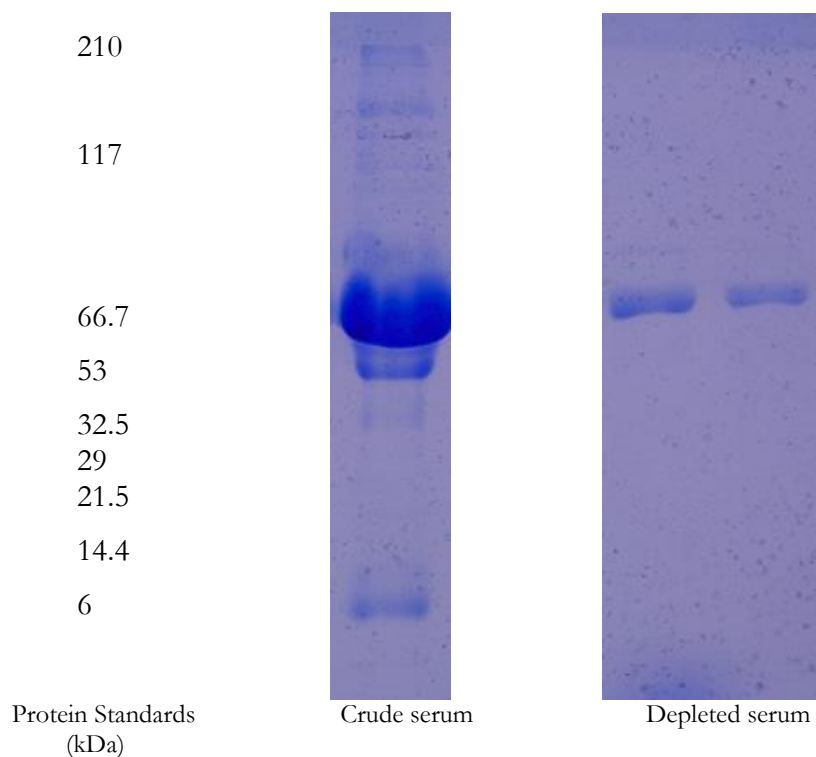


Figure 4.2 SDS PAGE of crude serum and depleted serum.

A review of the literature revealed that there were no studies into stool using SELDI-TOF technology and that there was little done in the field of proteomics using this sample type. Advice was sought from the manufacturer (Ciphergen) who had no experience with stool experiments. Conditions and chip surfaces for optimisation were selected. Four different chip surfaces were used to perform the optimisation, the H50 chip (strongly hydrophobic), a strong and weak cation exchanger CM10 and Q10 and a metal anion immobilized binding site chip the IMAC 30 were used with to perform the optimisation experiments. The samples were tested individually on each chip surface at a range of compatible pH's and the profiles inspected for quality and number of peaks. The H50 chip gave the most detailed profiles over a broad range of mass to charge ratios and this was selected as the optimum surface.

4.4 Analysis of 10 cancer serum samples compared to 10 normal serum samples on the H50 chip optimised for the low molecular weight region

The experiment was initially conducted using 10 cancer and 10 normal samples. The ProteinChip software data analysis package was employed for data analysis. The data generated from each chip was imported into one experimental file and the baseline was subtracted from each spectrum following which all data were normalized with a total ion current between 2000 to 20000Da using a calculated normalization factor of 1.99. All peaks were then detected using the automatic peak detection facility, a ratio of 10 signal to noise was selected with a minimum peak threshold of 30% of all spectra. There were 72 peaks greater than 2kDa, identified by the sample group statistics as being different between the two groups, of these there were 9 statistically significant peaks. The standard deviations were assessed for each value and those which had the largest standard deviations were identified for both cancer and normal groups. There was a subset of four biomarkers that appeared to be particularly significant, had small standard deviations relative to their value and separated the data out confidently. These four markers were at 3364Da, 3434Da, 3479Da and 3700Da. The mean peak intensity for each group and their standard deviations is shown in table 4.1 and the spectra are shown in Figure 4.3.

M/Z	P- VALUE	MEAN – CANCER	SD - CANCER	MEAN - NORMAL	SD - NORMAL
3479.1	<0.001	13	19	1.3	1.2
3364.6	<0.001	4.5	4.2	1.3	0.8
3434.6	<0.001	18	23	2.1	1.7
3700.8	<0.001	12	18	1.9	2.7
2841.9	0.001	0.1	0.3	1.1	2.2
5223.5	0.002	1.2	0.7	0.7	1.1
2066.4	0.029	1.0	2.5	0.1	0.5
3567.8	0.029	0.4	0.8	1.4	1.9
5782.5	0.035	1.5	1.6	0.5	0.4

Table 4.1 Discriminant peaks that distinguished between cancer and normal samples. P values, mean peak intensities and standard deviations. Highly discriminatory peaks are shown in light grey

The other markers identified had very low peak intensities compared to the background noise and so were not evaluated. The samples were then pooled and analysed, and the profiles shown below in Figure 4.4. The results showed that there was an almost total absence of the four peaks in the pooled normal sample although not a significant increase in the abundance of these markers in the pooled cancer sample as one might have expected.

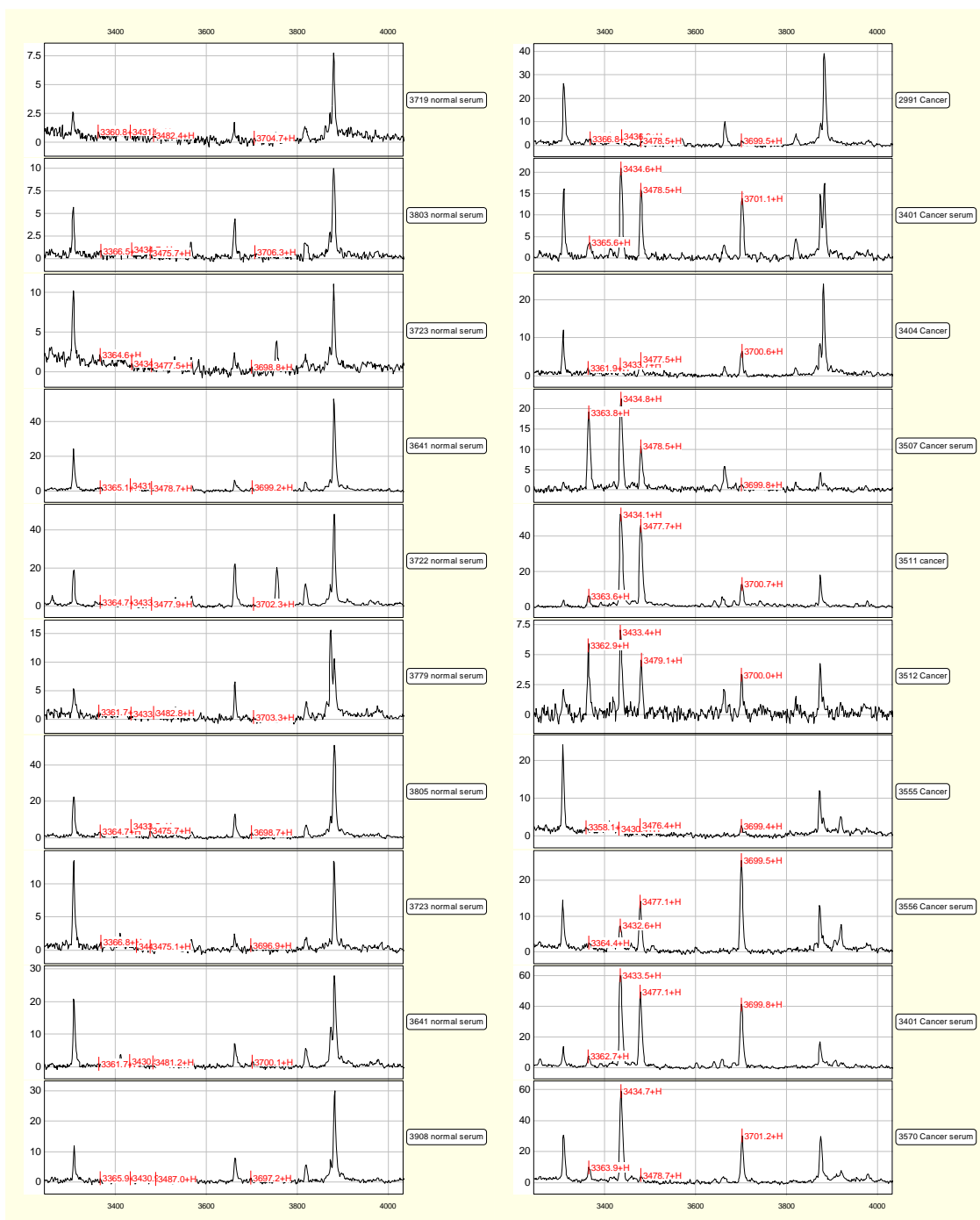


Figure 4.3

This shows the peaks detected in the 3kDa to 4kDa mass range that were identified as statistically significant between the two groups. The peaks at 3364, 3434, 3476 and 3700 are labelled in each sample. A clear difference can be seen between cancer and normal with a cluster of peaks that are absent on the normal samples are seen in most of the cancer samples.

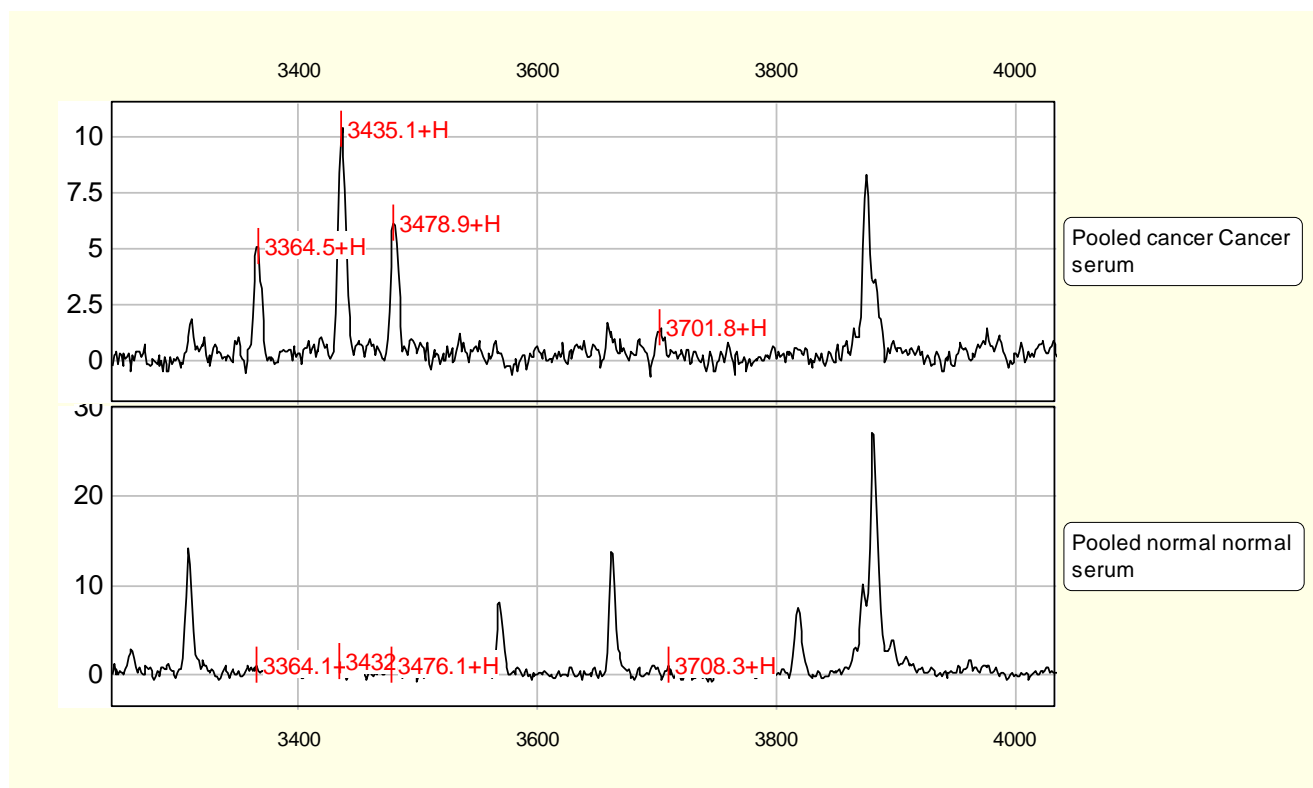


Figure 4.4
Pooled cancer and pooled normal serum samples over the mass range of interest. The peaks are labelled on each sample.

A log of the normalized intensity of the peaks between 3360Da and 3500Da is shown in Figure 4.5 and shows a separation of the cancer group from normal.

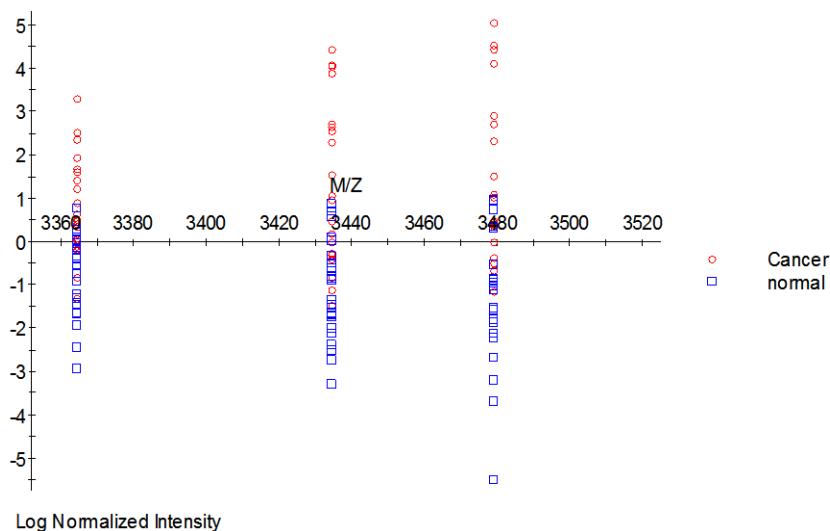


Figure 4.5 Biomarker wizard plot that shows the logarithm of the normalised intensity of three of the peaks that separate cancer and normal. 3.36kDa to 3.52kDa range.

The peaks were then used to identify cancer from normal as a screening tool, individually and together. Individually the peaks were not a powerful screening tool, the weakest peak 3479Da identified just 6 out of the 10 cancers. 3364 Da and 3434Da identified 8 out of 10 cancers and 7 of the 10 were identified using the 3700 Da peak. However, when combined the sensitivity was improved, with all 4 peaks identifying 9 out of the 10 cancers, a sensitivity of 90%. Table 4.2 shows the cancer detection rate of the peaks individually and combined. When the normal sample spectras were inspected for the same peaks, there were no false positives identified using the 4 peaks, a 100% specificity. One of the cancer samples that did not have the typical peak pattern displayed was a node positive cancer.

Peak	Number of cancers detected	Sensitivity
3364	8 out of 10	80%
3434	8 out of 10	80%
3479	6 out of 10	60%
3700	7 out of 10	70%
3434 and 3700	9 out of 10	90%
3364 and 3434	7 out of 10	70%
3364 and 3434 and 3479	7 out of 10	70%
3364 and 3434 and 3700	9 out of 10	90%
All 4 peaks	9 out of 10	90%

Table 4.2 Power of individual and combined peaks at identifying the cancer samples.

These results were very encouraging and the same peaks were expected to be shown on analysis of the full set of samples.

4.5 Analysis of 92 controls compared to 86 cancers using the H50 chip

The experiment was repeated this time with 86 cancers and 92 normal controls using the same conditions and chip surface. The same criteria were used for signal noise and stringency of peak detection. There were 8 samples that following the experiment were unsuitable for inclusion in the analysis due to error in technique and these were discarded from statistical analysis. Table 4.3 shows the peaks that were detected as being different between the two groups, peaks less than 2kDa have been discarded. There were 5 peaks that were different between the two groups that were significant, these are highlighted in bold and were not the same peaks that had separated the cancer and normal patients in the initial group of samples. 7901 Da, 8124 Da, 8566 Da, 8799 Da and 17 409 Da were the peaks that were of interest.

M/Z	P value	Mean – cancer	SD - cancer	Mean – normal	SD – normal
7901	0.00	5.67	5.99	7.37	4.87
8566	0.01	19.37	20.59	26.32	21.85
8124	0.01	10.97	11.02	12.59	8.78
17409	0.01	4.75	8.08	4.54	5.87
8799	0.03	20.51	19.16	24.00	15.90
33592	0.05	1.37	2.78	1.44	2.09
8669	0.06	19.55	19.15	21.25	14.27
14181	0.11	3.00	4.18	3.07	3.70
28303	0.14	7.64	13.15	7.45	12.22
45163	0.15	0.21	0.26	0.24	0.26
7758	0.17	18.16	23.01	21.85	22.35
14689	0.21	2.61	3.32	2.18	1.24
10255	0.40	2.91	4.28	2.58	2.98
28139	0.43	4.51	8.61	3.93	7.38
9303	0.48	11.94	12.35	12.95	13.79
8921	0.62	42.36	27.59	40.49	23.58
2342	0.67	4.66	6.89	4.71	5.59
3308	0.72	5.04	8.03	5.06	6.27
6827	0.81	4.86	7.64	4.51	5.42
6623	0.87	22.13	28.83	23.53	25.04
9143	0.88	15.66	14.26	14.53	12.61
4456	0.99	8.14	14.24	6.37	10.86

Table 4.3 Peaks that were detected as being different between the cancer and normal samples on analysis of the 100 normal compared to the cancer.

In this experiment the peaks that were detected were different from the previous experiment of ten samples. Additionally there was a change in the amount of the peaks detected rather than an absolute presence or absence. This finding is reflected in the logarithms of the normalised intensity (Figures 4.6 through to 4.9) that do not show much separation between the groups around the X axis. Instead there was an up or down regulation in the presence of these peaks.

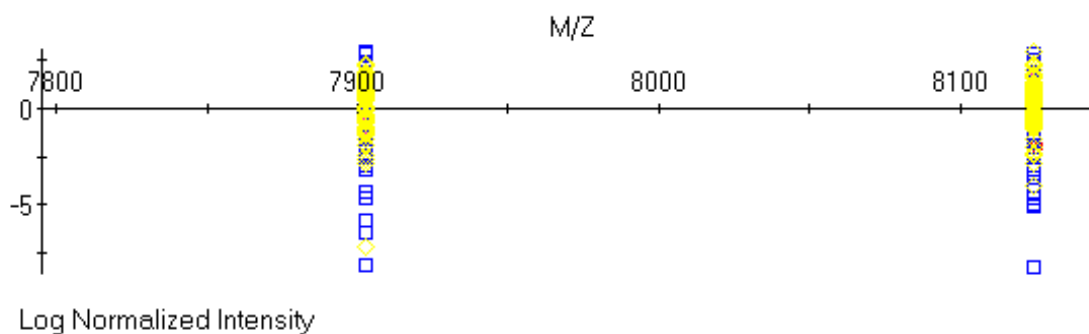


Figure 4.6 Logarithm of the peaks identified at 7900Da and 8124 Da as being significantly different between the two groups. MPI for cancer was 5.67 compared to 7.37 for the 7900 peak. MPI for cancer was 10.97 and 12.5 for normal for the peak at 8124Da. Cancer is the yellow diamond and normal the blue square.

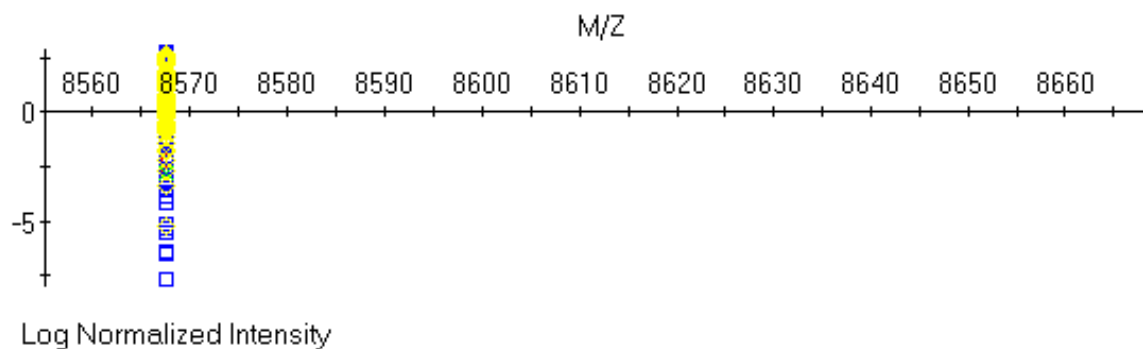


Figure 4.7 Logarithm of the peak identified at 8566Da as being significantly different between the two groups. MPI for cancer was 19.37 compared to 26.32 for normal.

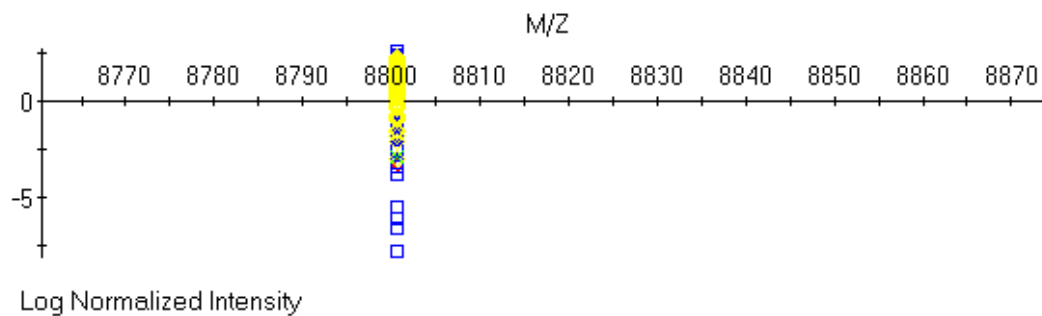


Figure 4.8 Logarithm of the peaks identified 8799Da. MPI of 20.5 cancer compared to 24 for normal.

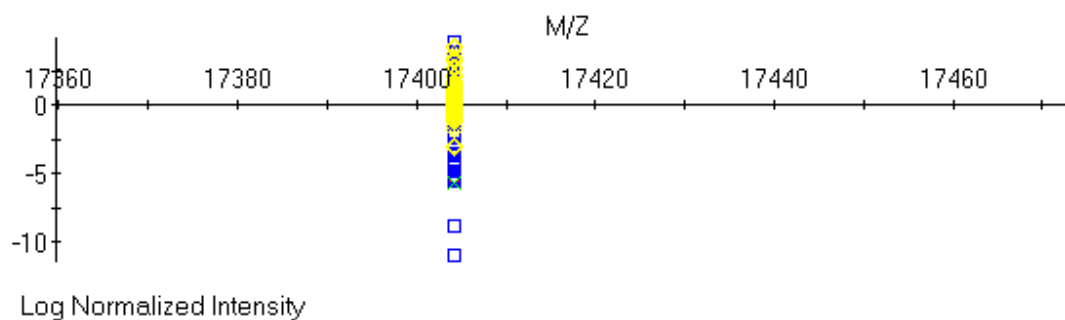


Figure 4.9 Logarithm of the peak intensities identified at the 17409Da peak. The MPI of cancer was 4.7 and 4.5 for the normal group.

The experiment was repeated and the profiles individually inspected for the four peaks that had so confidently separated out from cancer from the controls. The peaks were not present in significant numbers of samples and the peaks identified in the experiments were not reproducible between experiments. In order to reduce sample complexity it was decided to deplete the samples of the major serum proteins to unmask potential biomarkers that were not being bound due to competition from the more abundant molecules.

4.6 Analysis of serum on the H50 ProteinChip following serum depletion

Crude serum was taken and fractionated to remove the high abundance, large proteins and this was analysed using the H50 ProteinChip. The sample was split into Fraction1, Fraction 2 and a third fraction, the eluted bound proteins. The SELDI spectra for the high molecular weight range (50000Da to 200000Da) can be seen in Figure 4.10.

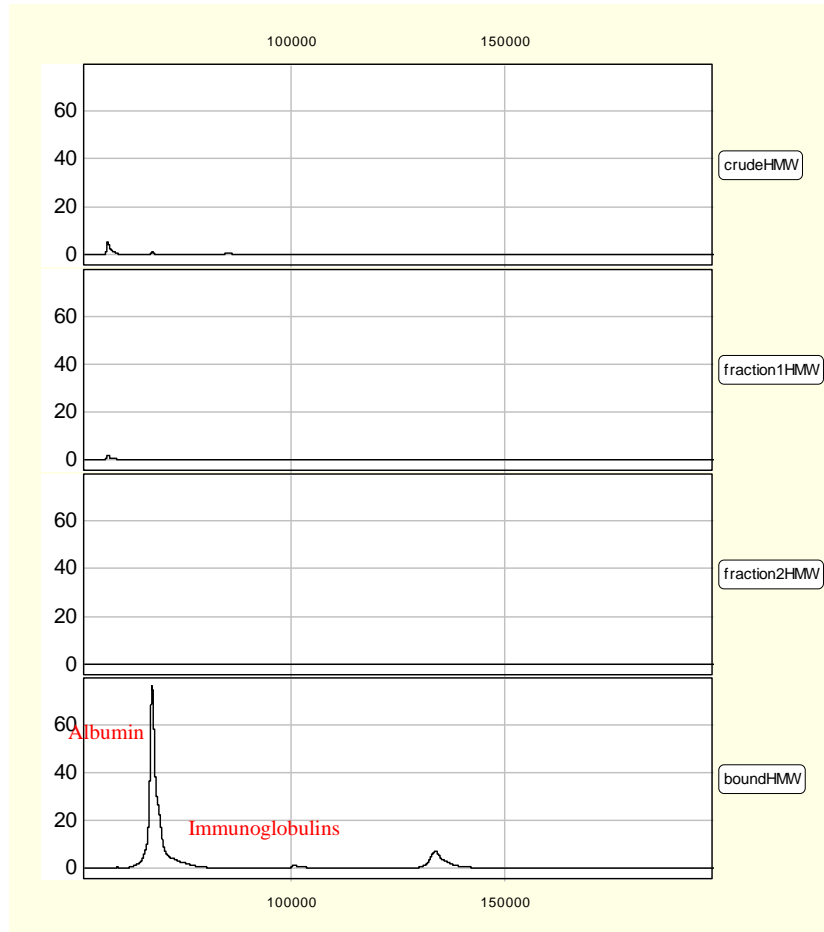


Figure 4.10 High molecular weight SELDI profiles of crude serum, fraction 1 depleted serum, fraction 2 depleted serum and eluted bound proteins. In crude serum there are few peaks. In the same sample the first eluted fraction is much as crude serums, on the second pass through the spin column the peaks are absent. When the proteins are eluted from the binding beads and the sample re-analysed the peaks are obviously seen.

Figure 4.10 shows that albumin (66000Da) is present in large quantities on the bound eluted protein fraction. Transferrin (90000Da), IgG (150000Da) and IgA (150000Da) are also not seen to be bound to the chip when incubated as crude serum and yet do bind when they are incubated as an eluted fraction. There are large amounts of each protein in the bound group,

reflected by their large peak intensities compared to the crude sample group. The peak profiles for the fractions that had been depleted of the large proteins were not detailed over the high mass range and there was little binding of proteins of this size.

The SELDI spectra for the low molecular weight range (2000Da to 70000Da) can be seen in Figure 4.11. Fraction one and the crude serum are very similar in their peak profiles, following the second run to create fraction two the MPI of each peak decreases significantly and there is not a corresponding rise in the number of peaks as was hypothesised. The depletion of serum revealed that rather than more detailed spectra being uncovered by this technique, the reverse was true and that depletion of the large proteins resulted in almost an absence of peaks on SELDI analysis.

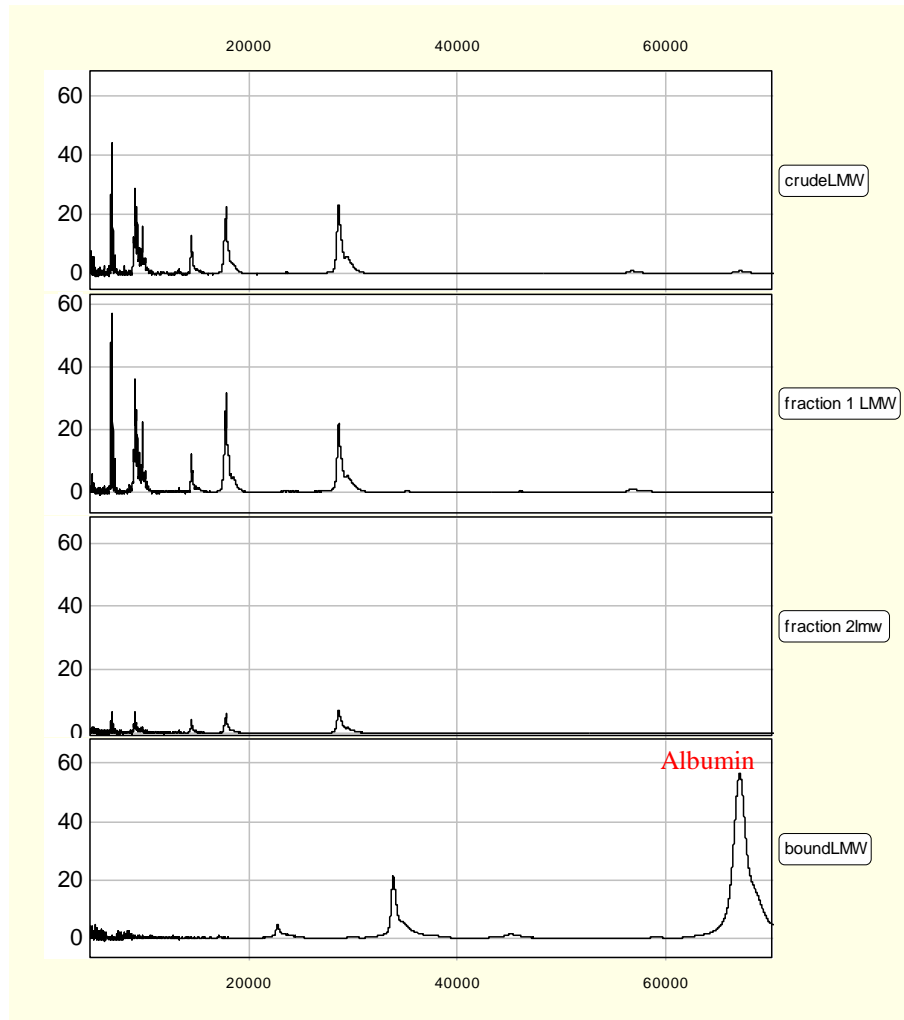


Figure 4.11 Low molecular weight SELDI profiles of crude serum, fraction 1 depleted serum, fraction 2 depleted serum and eluted bound proteins

4.7 Comparison of peaks in serum with those found in tumour

The peaks that had been detected in the tumour SELDI analysis that had proved of most interest were compared to those found in the serum (Table 4.5). Only one peak was within 10 Daltons of a peak that had been detected in each sample type, 3444 in solid tumour and 3434 in serum.

SOURCE	MASS TO CHARGE RATIO (DA)
Tumour derived peaks	2826, 3374, 3444 , 3490, 5497, 5686, 5717, 10854
Serum sample peaks	3364, 3434 , 3479, 3700 7901, 8566, 8124, 8799, 17409

Table 4.5 Peaks detected in tumour compared to those that were seen in serum.

CHAPTER 5 ANALYSIS OF HUMAN FAECES

SELDI-TOF analysis of cancer versus normal5.1 Introduction

5.1 Introduction

Composition of stool

Stools are accessible and despite the huge variety in diet and lifestyle between individuals they may represent a source of new colorectal cancer screening biomarkers. The content of an average 750g stool is 75% water, the rest being made up of undigested fibre and cellulose and other roughage with a rich mix of salts, mucus, bile, stercobilinogen and cells; both endogenous and bacterial. The large intestine is a site of intense protein turnover, on average 12g of protein rich material (0.5g to 4g total nitrogen) enters the large intestine each day mainly in the form of protein and peptides (20 - 30%). Dietary sources account for about 50% of this protein material, the rest is in the form of endogenous proteins; pancreatic enzymes, sloughed off cells and mucus(Hughes, Magee et al. 2000).

Stool is a complex and rich source of compounds and proteins and has not until now been examined using SELDI-TOF MS. Utilising this novel technology a method of examining stool was developed and this was then employed to analyse the stools of 47 cancer patients and 100 control subjects with normal colons.

5.2 Results

Samples were collected from the two groups, cancer (n=47) and normal (n=100). These two groups were then subdivided depending on whether or not the original sample had tested positive for occult blood. In total there were 47 cancer samples collected and 45 of these were found on rehydration and FOB testing to be FOB positive. This gave four sample groups, positive cancer (n=45), negative cancer (n=2), positive normal (n=22) and negative normal (n=78).

Measurement of protein concentration in stool samples

The stool samples used ranged in weight from 5g to 45g. Measurement of the protein concentration was carried out on five samples in each group using a Bradford assay and was found to be an average of 0.8mg/ml for the normal samples and 0.35mg/ml for the cancer samples.

Sensitivity of the Haemoccult II test

There were 47 cancer samples collected and of these patients 45 of them (95.74%) tested FOB positive. This gives an estimated sensitivity for the FOB as 96%. The number of positive FOB tests in the control group was 22 out of 100.

5.3 SELDI-TOF analysis of stool

A technique of analysis was developed after reviewing traditional proteomic methods of faecal protein identification. It was first necessary to establish that SELDI-TOF MS analysis of stool was possible and samples that were obtained from normal patients were used in the initial development of stool SELDI analysis. Despite the protein concentration being below the recommended levels, initial experiments showed that profiles could be achieved using faecal fluid samples from stool, optimisation experiments were carried out using stool from 3 normal samples (FOB negative) and 3 cancer samples (FOB positive) on different chips surfaces and using different binding buffers. The IMAC 30 chip and the CM10 and Q10 chips were selected at varying pH concentrations (pH4, pH5, pH6 for the CM10 chip and pH 7, pH8, pH9 for the Q10) (Figure 5.1). The number of peaks present in the profiles, as well as any differences between the peaks in the two groups were noted. Additionally the profiles were visually assessed for quality. Following this the optimum conditions for expression difference profiling were selected. The CM10 chip using buffers at pH4 was selected as the optimal profiling surface as this gave the best spectra. The quality of the profiles achieved is less than what is seen in serum analysis, with fewer peaks, more background noise and broader and less well defined peaks.

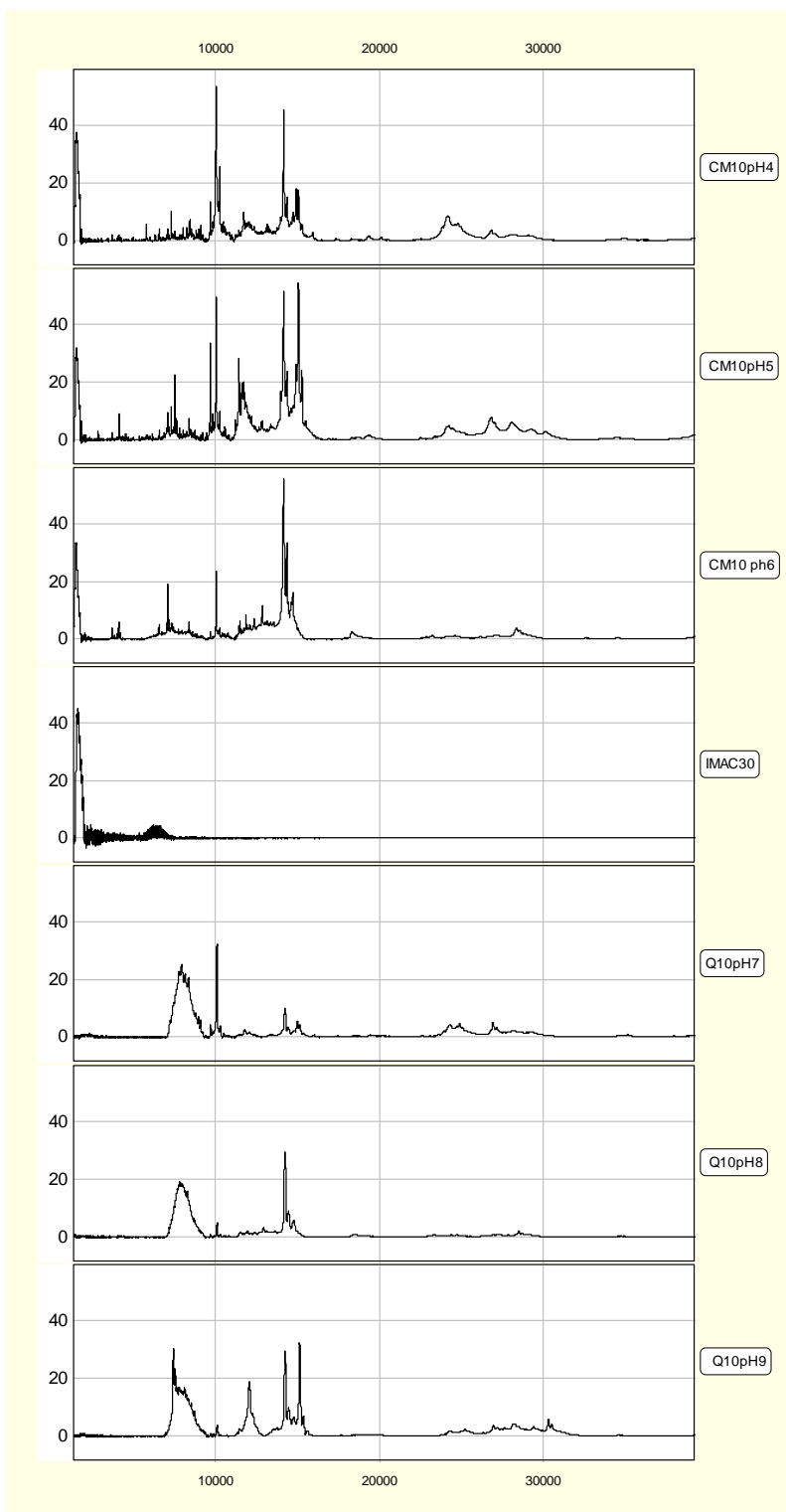


Figure 5.1 Chip surface and buffer conditions used for optimisation of stool method. The profiles obtained on the CM10 chip at pH4 were shown to be most detailed.

Seven experiments were then designed and carried out on the CM10 chip at pH4 using the same spot protocols each time. The number of FOB negative cancer samples was small (n=2).

- Cancer (n=47) compared to Normal (n=100)
- FOB positive (n=67) compared to FOB negative (n=80)
- FOB positive cancer (n=45) compared to FOB positive normal (n=22)
- FOB positive cancer (n=45) compared to FOB negative normal (n=78)
- FOB negative cancer (n=2) compared to FOB negative normal (n=78)
- FOB negative cancer (n=2) compared to FOB positive normal (n=22)
- FOB positive cancer (n=45) compared to FOB negative cancer (n=2)

5.3.1 Comparison of SELDI stool profiles between cancer and non cancer controls

Two bioprocessors were used to analyse the 147 samples, using internal calibration. The profiles had their base line subtracted and were visually inspected for quality and outliers and then normalised. There were 11 normal profiles removed from the analysis as they were not of sufficient quality. Highly stringent selection criteria were used, peaks selected were present in 50% of profiles. There were 76 peaks identified between the two groups, 6 were less than 2kDa and so were not considered in further analysis. There were 11 peaks that were significantly different between the two groups, they are shown in the table 5.1.

MASS TO CHARGE RATIO (DA)	P VALUE	MPI NORMAL	MPI CANCER
3437	0.043	13.4	4.8
3521	0.043	5.6	0.6
3574	0.043	6.1	0.0
3973	0.021	9.0	2.5
4633	0.021	3.1	19.4
9801	0.021	5.0	0.0
10838	0.021	7.6	1.4
16511	0.043	1.3	5.9
33423	0.021	0.1	0.7
37087	0.021	0.0	0.4
47026	0.043	0.0	0.5

Table 5.1. This shows the peaks detected by biomarker wizard that were significantly different between normal control stool samples and cancer stool samples. The peaks highlighted in bold were up regulated in the cancer samples.

The peaks at 16511Da and 33423Da were confirmed to be the same molecule occurring as a dimer. Each peak was then inspected across the spectra and visually assessed. Dimer peaks were identified and removed, this is done by a software tool that is built into Ciphergen that can compare two peaks and identify if they are dimers; a peak at 60000Da and a peak at 120000Da can represent two albumin proteins stuck together. There were 6 peaks that had higher values in normal stool sample compared to the cancer stool samples and 5 peaks where there was an abundance in the cancer stools. In particular, there was a peak at 4663Da which was considerably higher in the cancer group than in the normal (MPI 3.1 versus 19.4). The peaks that were increased in the stool of the cancer patients were then compared to the known peaks that had been detected in the tumour samples but there were no peaks that had been detected in the tumour profiling experiment that were detected in the comparison of cancer versus normal (Table 5.2). The peaks that were upregulated were then used as a

screening tool, using all the peaks and also individually to differentiate the cancer samples from the controls. The dimer peak of 16511 and 33423 were treated as one peak.

PEAK	NUMBER OF CANCERS DETECTED	SENSITIVITY
4633	37 out of 47	78%
16511	27 out of 47	57%
37087	32 out of 47	68%
47026	16 out of 47	34%
4633 and 16511	37 out of 47	78%
4633 and 37087	37 out of 47	78%
4633 and 47026	37 out of 47	78%
4633 and 47026 and 37087	39 out of 47	83%
All 4 peaks	39 out of 47	83%

Table 5.2 The power of individual and combined peaks that identify cancer samples from normal.

5.3.2 Comparison of SELDI stool profiles between FOB positive and FOB negative patients

Two bioprocessors were used to process the FOB positive stool samples (n=67) and the FOB negative stool samples (n=80) with internal calibration. The spot protocols used were the same as for the cancer and normal experiment. The difference between an FOB positive profile and an FOB negative profile was obvious and is shown in the gel view of two samples in Figure 5.2.

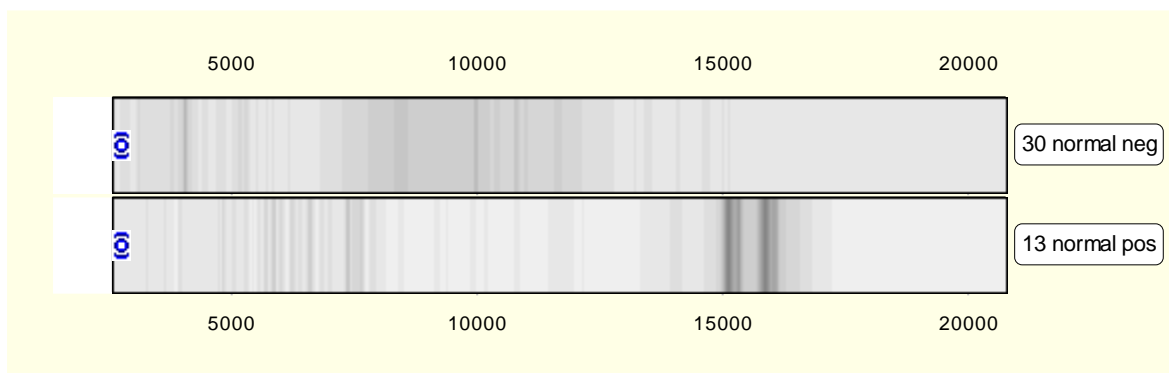


Figure 5.2 Gel view of an FOB positive trace versus an FOB negative trace. There are two dark and distinct peaks visible around the 15kDa region. These peaks are in keeping with the α and β chains of haemoglobin.

There were 37 peaks identified by the biomarker wizard using the same 50% peak presence rates. At the less than 2kDa range 5 were discarded as being due to matrix noise, only 4 were statistically significant between the two groups. (Table 5.3)

MASS TO CHARGE RATIO (DA)	P VALUE	MPI NORMAL	MPI CANCER
2228	0.02	1.9	4.2
15152	0.01	6.2	14.7
15894	0.01	1.0	12.3
29287	0.04	0.6	0.1

Table 5.3 This shows the peaks detected by biomarker wizard that were significantly different between FOB positive samples FOB negative stool samples. The peaks highlighted in bold were up regulated in the cancer samples.

These peaks were compared to the peaks that were identified in the tumour experiment and there was no correlation between these peaks and the ones expressed from the tumour.

Analysis of the data revealed that there were 7 FOB negative patients that had small peaks at the 15kDa range. These samples that had tested as FOB negative but appeared to have a profile similar to the FOB positive samples and so were retested in a separate experiment to eliminate the possibility of contamination. SELDI analysis confirmed the initial findings and

the peaks persisted at 15kDa, they are clearly identifiable on the spectra. This eliminated the possibility of contamination of the samples during the initial comparison. Figure 5.3.

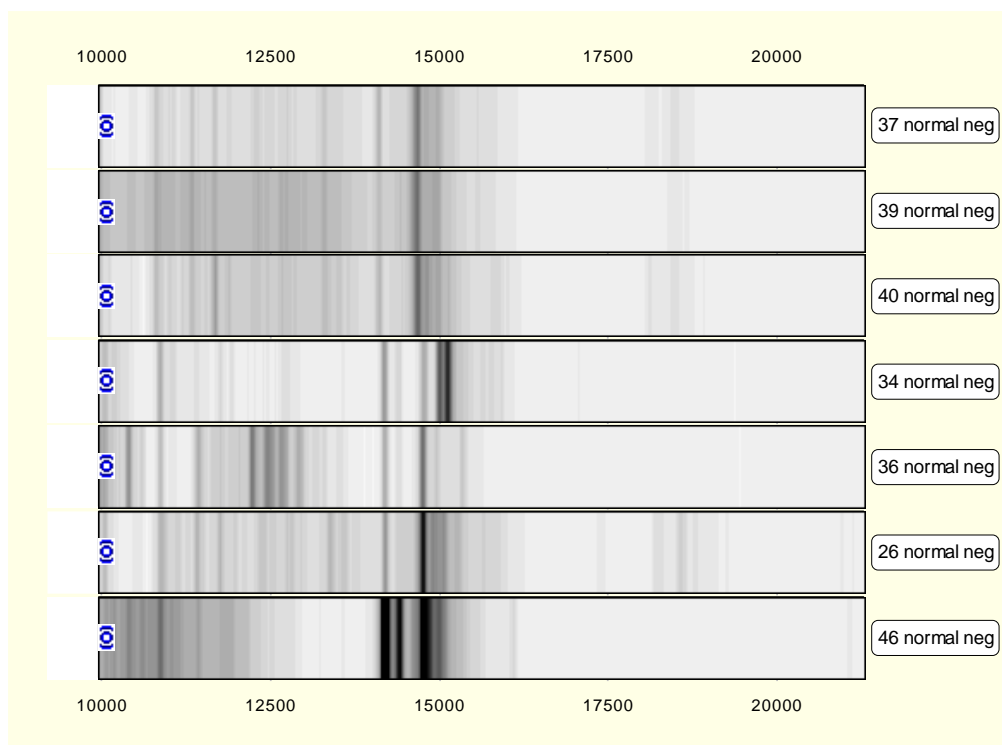


Figure 5.3 Gel view of the re analysis of the FOB negative normal samples that have peaks at 15kDa

5.3.3 Comparison of FOB positive cancer (n=45) and FOB negative normal (n=78) samples

In this experiment two bioprocessors were employed with internal calibration to compare the FOB negative normals and the FOB positive cancer groups. These were compared using the Biomarker wizard and a panel of 6 significant peaks were identified. Peaks at 5135Da, 6334Da, 7929.5Da, 15130 Da, 15879Da and 66010Da were significantly different between the cancer positives and the normal negatives these are shown in table 5.4. Comparison was made with the peaks identified in the tumour experiment and there was no correlation between these peaks found to be of significance separating FOB positive cancer and FOB negative normal and the ones expressed from the tumour.

MASS TO CHARGE RATIO	P VALUE	MPI +CANCER	MPI -NORMAL
5135	0.01	24	2.7
6334	0.02	55	25.1
7929	0.02	12	3.34
15130	0.03	9.6	1
15879	0.03	13	3.03
66010	0.03	16	3.97

Table 5.4 Statistically significant peaks that differentiate between cancer FOB positive and normal FOB negative samples that were upregulated in the stool of cancer patients. The mean peak intensities are shown for each.

5.3.4 Comparison of FOB negative cancer (n=2) and FOB positive normal (n=22)

There were only 2 FOB negative cancer samples in this study which limited the application of statistical analysis of these groups. Using one Bioprocessor and internal calibration the samples were analysed. Using values greater than the mean peak intensities for FOB positive cancer at 4633, 16511, 33423, 37987, 47026 Da and values less than the mean peak intensities of the down regulated peaks in all cancer at 3437, 3521, 3574, 3973, 9801 and 10838Da it was possible to identify the two cancer samples as belonging in this group.

5.3.5 FOBt positive cancer (n=45) and FOBt negative cancer (n=2)

The limitations in this comparison again related to the small number of FOBt negative cancer samples. Analysis of the samples of the FOBt positive cancers using the peaks previously identified in the FOBt positive samples compared to the FOBt negative samples was undertake and these peaks (5135, 6334, 7929, 15130, 15879, 66010) were not seen in the FOBt negative cancers.

5.3.6 FOB negative cancer (n=2) and FOB negative normal (n=78)

This experiment compared the 2 FOBt negative cancers with the FOBt negative normals. Due to the difference in sample sizes, the peaks that had previously been found to distinguish cancer from normal were used as a screening tool. Using this panel of peaks it was not possible to identify the two cancer samples as belonging in this group.

5.3.7.FOB negative normals (n=22) and FOB positive normals (n=78)

FOB positive normals and FOB negative normals were not compared in one experiment, therefore the peak profiles were taken from two experiments to allow comparison of these groups. This is a less valid method of comparison as results are affected by processing in separate experiments. However, the peaks that were significantly different between the groups were likely blood proteins, all around the 15kDa size. This finding is difficult to interpret without repeating an experiment to allow a direct comparison. The table below shows the MPI for the negative and positive groups and the standard deviations. There was less standard deviation in the negative FOB samples, although there were peaks that were FOB negative on testing that showed small peaks at 15kDa suggesting a sub-FOB detectable positivity that was seen by the mass spectrometer.

M/Z	p	Mean - negative	SD – negative	Mean - positive	SD - positive
14088	0.01	1.79	1.48	7.99	4.21
14711	0.01	2.27	0.35	7.17	5.34
15147	0.01	6.56	3.57	31.89	4.77
15344	0.01	2.32	1.96	21.91	6.01
15554	0.01	0.86	1.27	9.72	2.35
15894	0.01	5.76	2.78	30.78	6.14

Table 5.5 Significant peaks that were detected between FOB positive normals and FOB negative normals. The peaks are all around 15kDa, as the profiles have been taken from 2 experiments it is difficult to conclude much, the variation around the 15kDa point may represent this variability.

5.3.8 Samples may have degraded over time following storage

In this experiment samples that were collected in the initial optimisation were defrosted and incubated on a chip surface under the same conditions as their initial analysis. The amount of detail in the spectra in the samples older than one year was seen to be less than that of the younger samples and this is shown in Figure 5.6. As samples were stored at -80 C in a controlled environment, degradation seems an unlikely source of the variability, and this may represent the change in results with SELDI profiling over time, one of the criticisms of this technique being intra-laboratory as well as inter-laboratory variability.

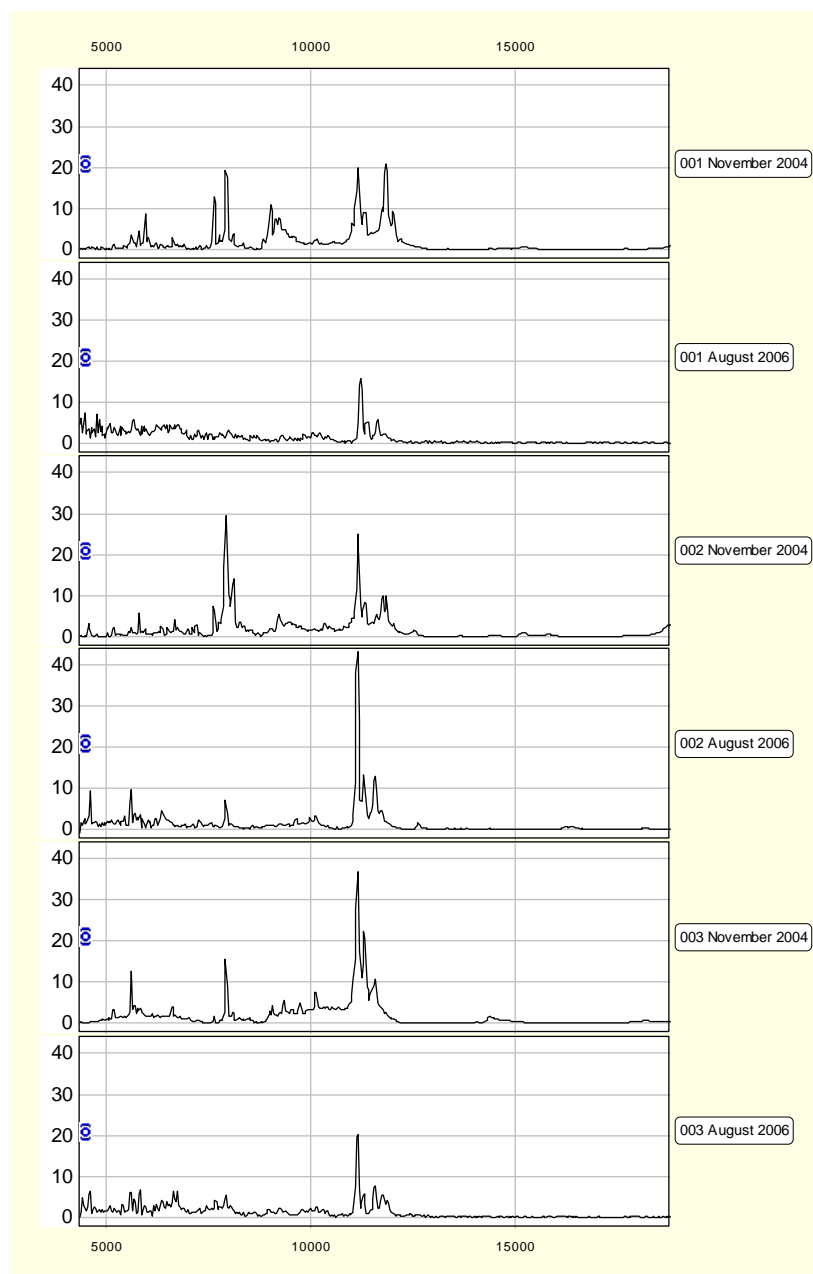


Figure 5.6 SELDI-TOF spectra achieved are altered following storage of 21 months at -80

5.4 Comparison of peaks in stool with those found in tumour

The peaks that had been detected in the tumour SELDI analysis that had proved of most interest were compared to those found in stool (Table 5.5). The peak at 10838 in the stool samples was down regulated in cancer and a similar sized peak at 10854 is found to be present in tumour.

SOURCE	MASS TO CHARGE RATIO (DA)
Tumour derived peaks	2826, 3374, 3444, 3490, 5497, 5686, 5717, 10854
Stool sample peaks	Up regulated in all cancers: 4633, 16511, 33423, 37087, 47026 Down regulated in all cancers: 3437, 3521, 3574, 3973, 9801, 10838 Upregulated in all FOBt positive cancers: 5135, 6334, 7929, 15130, 15879, 66010

Table 5.6 Peaks detected in tumour compared to those detected in stool samples.

CHAPTER 6

DISCUSSION

This research was undertaken from 2006 to 2008 using what was at that time, a truly novel and promising technology. A landmark paper in the Lancet in 2002 (Petricoin, Ardekani et al. 2002) had shown that using SELDI they were able to discriminate between ovarian cancer and other benign gynaecological disease with 100% sensitivity and a 98% specificity. The resulting media attention resulted in a US congress resolution urging more research in this field. This paper was then widely criticised by many traditional mass spectromitists and was subject to a lot of debate but in spite of this many people began using this technology as an avenue of biomarker discovery. It was promoted as a high throughput and high sample complexity method of proteomics compared to traditional methods of multi dimensional gels which could analyse only one sample at a time. Over the last decade, the field of proteomics has moved on and SELDI-TOF technology has not heralded a new era of personalised medicine and easy to discover highly specific and sensitive tests for cancer and other diseases. More sophisticated methods of proteomic analysis have overtaken it, allowing quantitative and more reproducible results to be obtained. SELDI may have a role in looking at evolving changes in the proteome of patients, for example in response to treatments such as chemotherapy.

Pre-analytical phase standardization has been found to be crucial in achieving good results in proteomics studies as well as collection and handling of the samples following sampling. Proteomics is now known to consist of many subspecialty fields, cytoproteomics the study of the intracellular proteome and serum subproteomics, the study of circulating

microparticles. It is a far broader and more complex field than perhaps was fully grasped twenty and even ten years ago.

Many high profile institutions, for example MD Anderson in Texas have published papers retracting work previously published using SELDI-TOF, the perceived differences in cancer and non-cancer samples were on re-analysis traceable back to the day that the experiments were run, significant results being achieved by varying the day of the experiment. Increasingly advanced statistical packages are being used to mine data sets.

The key in SELDI-TOF and all proteomics studies remains that the peaks must be identified and evaluated as a biomarker following discovery.

6.1 Study Design and Sample Collection

Funding for this work had been secured prior to commencing the project; the initial 4 months were taken up with reviewing the literature which was carried out in 2 parts, a review of screening for CRC and SELDI-TOF as a platform for biomarker discovery.

After becoming competent at basic laboratory techniques (tissue culture, protein quantification, gel preparation, making stock solutions) and then attending a training course run by Ciphergen in order to use the machine, the main body of work was begun. During this time ethical approval was sought. Sample collection could then begin and work on tumour analysis and a method of analysing stool samples was developed.

Initially we hypothesised that the proteins that we might detect in the stool or in the serum may perhaps be present in the tumours themselves. Also, this was a way to gain experience with the technique and how to optimise conditions for analysis and use the statistics.

This revealed markedly different spectra between cancer and normal and following further review of the literature, it was decided to try and combine LCM with SELDI as this was a purer sample as there is no contamination by connective tissue and other cells. The literature has many studies comparing 5 tumours with 5 areas of normal tissue, the reason for this being that you need at least 2000 individually dissected cells to achieve an adequate amount of protein to get just usable spectra. This was prohibitive from a time point of view as to dissect 2000 cells for one spot takes many hours of work.

In designing this study sample collection and handling were considered carefully. When planning the number of patients and samples to include it was not possible to use a traditional power calculation as there is no available power calculation for SELDI profiling.

Sample size was justified based on the time available for sample collection and the number of cancers that were seen each week in Ninewells Hospital.

Stool sample collection proved one of the most difficult aspects of this project, collection was performed for the cancer group over a 10 month period and despite best efforts to recruit as many patients as possible, the initial anticipated number of 100 cancer samples was not achieved. Many of the cancer samples were collected pre-operatively when patients were undergoing bowel preparation. This resulted in a sample that was very high in water content and many of these contained too low levels of stool to give a usable profile and so were excluded from analysis. Other reasons for the failure to collect a larger number of stool samples included; samples wrongly being sent to the laboratories for microbiological testing for microbiological testing, patients forgetting to collect the samples and samples going missing from the ward.

The control stool samples were of higher quality in terms of the amount and were of normal consistency but as they were mailed in samples of fresh stool the age of these samples, what storage conditions and transport conditions were like for each was not known. None exhibited any macroscopic mould or bacterial overgrowth. Although this method of sample collection presented a greater degree of variation for each sample it is also the basis of the current screening programme and so was used as it was already known to be acceptable to patients. This time delay may account for some variation between the samples.

The control patients were recruited from the screening programme and so do represent a selected group of patients in that they were all recently FOB positive. Some patients

reported when they were consenting to take part in the study that they were reluctant to handle their bowel movements although this did not prevent anyone in the cancer group taking part in the study. The amount of stool returned by the control patients was in general very low, perhaps as only a small amount is required by the screening programme. In contrast the hospital acquired stool samples were usually larger, perhaps as these patients had been interviewed face to face and were undertaking bowel preparation.

It would have been challenging to collect fresh stool samples from each group, in particular normal controls and to immediately process and store them and this was considered to be beyond the resources available. The bowel preparation for colonoscopy within the screening programme is usually excellent and although taking the samples at this time was considered as a potential source of stool samples it would not have reflected a normal stool. The preparation of the samples for SELDI analysis was developed based on methods of stool preparation used for traditional proteomic faecal analysis as there was no manufacturer protocol or previously published method for this. Development of this was time consuming and initial methods of mixing stool in various buffers in much the same way as serum is handled were not successful at obtaining useable spectra.

The quality of the spectra obtained, in terms of the signal to noise ratio and the number of peaks seen and the amounts of each is less clean than the profiles achieved when examining serum and this has perhaps resulted in only the largest and most detailed abundant peaks being examined. Perhaps with techniques of sample fractionation or pre analysis processing, cleaner peaks could have been achieved. In particular, the samples could have been dehydrated and then rehydrated as a more concentrated solution of protein.

It was possible to better control the variables in the collection of serum, the control samples came from patients undergoing screening colonoscopies. They were all fasted pre procedure, had undergone the same bowel preparation and had proven normal colons. Fasting itself produces a metabolic response and this has been extensively studied in the medium and long terms; there may be changes in the serum proteome during the period that the colonoscopy group are fasting. Cancer serum samples were collected pre-operatively at pre-admission clinic or on the ward prior to theatre, some were fasted but most were not. This level of homogeneity within the control group is perhaps a disadvantage as it does not reflect the real life conditions that a screening test would need to overcome in order to detect disease.

6.2 SELDI-TOF MS analysis

Methods of sample handling, sample preparation for analysis and also how the data were handled following analysis in the mass spectrometer all introduced variables into the experimental design. Preparation of the array was crucial and the experiments described in this thesis were all done manually and not through an automated process. One of the disadvantages of manual handling of the samples was that processing of each sample was time consuming and so the conditions for each sample on each chip surface were slightly different as they may have incubated longer and be subject to more air drying and more time exposed to the matrix as well as different lengths of time spent on ice. Methods of processing the samples could also have impacted greatly on the results, the selected conditions used in the experiments described here were chosen after optimisation which identified the most detailed profiles. It is possible that this resulted in a biomarker that has not bound to this selected chip surface being overlooked. Additionally, in order to avoid bias the samples were not incubated in batches of cancer and normal but were instead spread across the Bioprocessor. Ideally, samples could have been blinded for analysis but this was felt to introduce the potential for error in sample handling.

The SELDI machine itself is a further source of variation and requires validation and calibration weekly in order to carry out experiments. Even with this, there is a day to day variation in the measurement of the mass charge ratios of standard, known calibrants. It has so far not proved possible for groups working in different laboratories to reproduce each others results when using SELDI detected peaks.

The data produced in a SELDI profile shows clear and detailed peak profiles for the smaller molecular weight proteins (below 20 000Da) but above this, even when the machine is optimised to show the larger proteins, there are far fewer peaks seen. This is not in keeping with the sizes of many known serum proteins which are generally larger and it seems that these proteins are being left out of analysis using this method of proteomics. Additionally, the abundance of proteins or protein fragments that are present on a profile do not necessarily translate to an abundance *in vivo* or even in the serum sample itself as this may represent instead the protein that binds most efficiently to a given chip at given conditions. However, SELDI should be considered as a platform for biomarker discovery and is not intended to be a diagnostic tool.

Differing methods exist for analysis of the data that is generated. There are two software packages that are marketed by the makers of the machine, one is the integral Biomarker Wizard Software that comes with the protein chip reader and provides univariate analysis and the other is Biomarker Patterns software which makes a decision tree based, multivariate analysis of the data in order to differentiate between two or more groups. Univariate analysis was used in these experiments, utilising the inbuilt biomarker wizard software. Multivariate analysis using many layers of decisions and presence or absence of peaks and up or down regulation is not likely to lead to development of a screening test as we currently know it and seems to be a move away from identification of a single disease biomarker. Furthermore a decision tree looking at up and down regulated peaks in SELDI profiles is not applicable as a screening tool due to the poor reproducibility of SELDI data when used outside of individual laboratories.

6.3 Analysis of Tumours

Using traditional methods of tumour homogenisation, usable and clean SELDI spectra were generated, and the differences between the two tissue types were marked with 50 different peaks detected. The large number of significantly different peaks between the two groups suggests that the samples were representative of their tissue type. LCM would have resulted in only cancer cells and normal cells being compared with none of the surrounding connective tissue and blood vessels. Tumour was analysed with LCM and subsequently analysed with SELDI-TOF but this proved too time consuming to collect a sufficient amount of material to reach the required protein concentration and was not suitable to use as a method within the time constraints of this project.

The peaks identified by the software that were different between the two tissue types were multiple as would be expected. Disappointingly none of the peaks that were present in the tumour samples were present in the serum samples either in the initial optimisation batch or in the larger experiment of all the samples. The same was also true with the stool samples, the peaks that were identified did not appear to be the same. Even allowing for intervariability over time with the machine there were not any candidate markers that were traceable through tumour, serum and stool.

6.4 Analysis of serum

SELDI-TOF technology has been applied here with variable success, the main problem being interassay variability, with peaks that are present in samples that are then absent or not as statistically significant on repeat testing. There was not any demonstrable intra-assay variability and the samples that were repeated in triplicate were consistently reproducible. This variability could be explained by errors in sample processing; performing the experiments on a large scale involved using three bioprocessors and this greatly affected the length of time for which each sample was incubated for and the complexity of the experiment. It is possible to exclude freeze thaw cycles as a source of error as the samples were stored in small aliquots following the initial freeze thaw cycle.

It is hard to ignore the results from the training set as the peaks that were found were significant. This pooled sample of serum could be further analysed to identify these peaks (using iTRAQ technology for example) and this work is ongoing. By identifying these peaks further research could be undertaken to evaluate this result.

The depletion of the major plasma proteins was employed to address this issue, as it was thought that by removing the biggest and most abundant proteins this would reveal a subproteome of peptides and proteins that were not managing to bind to the limited number of binding sites in each chip. Presumably due to the “stickiness” of albumin, depletion of the serum did not reveal lots of peaks and the profiles were almost blank.

Human error was initially considered as the reason for the differing results and the large experiment was repeated using aliquots of the same samples, this again failed to reveal the

peaks that were present in the training set. The collection and storage conditions for the samples were consistent and are unlikely to be significant in explaining the results.

Contamination during the experiment is possible but as the samples were not incubated on one cancer chip and one normal chip this seems unlikely.

Testing for CA242, CEA and CA19-9 was not carried out and it is possible that these known biomarkers may be present in the samples although there were no peaks seen at the molecular weight of these proteins; CEA measures 180,000Da and CA242 is also a very large protein of over 200 000 Da.

6.5 Analysis of faeces

Stool samples were collected in universal containers and stored at room temperature as any new screening test must be robust. It was not feasible to devise a method of sample collection that would allow immediate storage of samples at -80°C . The Hemocult gFOB test was employed in all stool samples on receipt, in the case of the cancer samples this was performed in the ward shortly after the time of collection. The samples were either tested prior to the addition of PBS and homogenisation or after when the sample was judged to be too dry to sufficiently smear onto the developing window of the testing kit. The control stool samples were at least 24 hours old, and likely even older. Only 22 of the samples were gFOBT positive (out of 100). It may be that the patients no longer had blood present in their stool samples, or it may be that the prolonged transit time has degraded the haemoglobin. The samples were collected from control patients up to 16 months following their screening colonoscopy.

The stools were homogenised on receipt and mixed with PBS to obtain a mixture suitable for analysis. It may have been more useful to collect the external part of the stool that had been in contact with the (distal) colonic mucosa as this may have had more contact with potential biomarkers. It is unlikely that this would have made any difference in detecting right sided cancers as the stool is liquid on entering the right colon.

To the best of our knowledge this is the first time SELDI-TOF MS has been applied to analyse human faeces. This novel technology has been successfully used here and may have implications for research in other conditions of the colon and rectum and may even be applicable in proximal gastrointestinal disease. Although not as detailed as the serum profiles, the quality of the faecal profiles was adequate to allow analysis. In spite of the

hypothesis that intracellular proteins may be actively leaked into the lumen of the bowel and be present in the stool samples, this was not reflected in the studies of stool.

It seems that blood proteins are present in the stool samples and are detected even in those samples that tested FOB negative using the Haemoccult, suggesting that SELDI is more sensitive than the gFOBt at detecting blood in stool samples. There were 6 peaks that were upregulated in the cancer stool samples and due to the low number of FOBt negative cancer samples (2) it is not possible to be sure that these are all blood proteins. The small amount of protein that was present in the stool samples prohibited attempts to identify these proteins using traditional methods of analysis and the kit available from Ciphergen Biosystems (Fremont CA) was also not recommended for faecal protein identification. It was anticipated that in the comparison and analysis of stool samples, the degree of variation within individuals' diet may lead to contamination with a large number of exogenous proteins from animal sources but the uniformity of the spectra does not support this. The spectra were not as detailed and as clear as the usual SELDI spectra and fewer peaks were seen in the stool spectra compared to the tumour and serum derived peaks.

Peaks that were present in more abundant quantities in the control samples are perhaps more difficult to explain but a potential source of this change is that some of the cancer samples were collected from patients undergoing bowel preparation and that these samples were perhaps more dilute because of this. Another possible explanation may be dietary factors.

6.6 Overall comparison of peaks

There was not one peak that was reproducibly seen in tumour, serum and stool. However, the variability of SELDI-TOF MS means that some peaks that have values close to one another may represent the same proteins, although the significance of these is unclear without formal identification. 3444 in the tumour samples was within the margin of error to be appearing in the serum as 3434, and there is also the possibility that some of these peaks are peptide fragments of larger proteins. Table 6.1 shows the profile peaks that were detected in each experiment as significant.

Source	Mass to charge ratio (Da)
Tumour derived peaks	2826, 3374, 3444 , 3490, 5497, 5686, 5717, 10854
Stool sample peaks	Up regulated in cancer 4633, 16511, 33423, 37087, 47026 Down regulated in cancer 3437, 3521, 3574, 3973, 9801, 10838
Serum sample peaks	3364, 3434 , 3479, 3700 7901, 8566, 8124, 8799, 17409

Table 6.1 Table showing the peaks identified in the tumours, stool samples and serum. The peaks highlighted in bold are similar in size.

It was hoped that there would be a peak present in the tumour that would be either secreted or leaked into the serum or stool and that this would be identified by SELDI-TOF analysis.

This was an ambitious and perhaps overly optimistic assumption to make.

CONCLUSIONS AND FUTURE DIRECTIONS

Presented in this thesis are several findings that represent future avenues of exploration in the hunt for a novel screening test for colorectal cancer. During the time that this period of research was undertaken, much has changed in our understanding of SELDI TOF MS and proteomic technology and the screening programme has reached full national coverage.

Analysis of adenocarcinoma and matched normal mucosa has been carried and there was a marked difference between normal mucosa and adenocarcinoma likely to represent highly abundant intracellular proteins. Interestingly, the peaks that have been uncovered here are similar in size and charge to those found by another group (Krieg, Fogt et al. 2004). In spite of the hypothesis that intracellular proteins may be actively leaked into the lumen of the bowel and be present in the stool samples, this was not detected in our studies of stool.

We are the first to have developed and described a method of achieving acceptable spectra using the SELDI technique. We have shown that blood proteins are present in the samples and that these peaks are detected even in those samples that tested FOB negative using the Haemoccult suggesting that SELDI is more sensitive than the FOB at detecting blood in stool samples. The profiles generated, although usable were not of high quality and are unlikely to lead to further areas of research using this technology.

Analysis of serum using SELDI has shown huge variability in the detected peaks. Samples tested in the initial optimisation on one chip surface present very different results when performed on a later date on a larger scale and this variability is problematic when

developing a screening test which must be robust. It is concerning that the results are variable to such a degree and it may be that the peaks detected do not represent whole proteins and are more likely to be fragments of proteins.

SELDI-TOF MS analysis is not applicable as a screening modality but may represent future avenues of biomarker discovery and combined with quantitative technologies which have less variability and allow for protein identification a new screening test may be found.

References

- Adam, B. L., Y. Qu, et al. (2002). "Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men." Cancer Res **62**(13): 3609-14.
- Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.
- Ahlquist DA, C. A., Jett J, Pearson RK, de Groen PC, Edell E, et al. (2000). "Universal detection of aerodigestive cancers by assay of nonapoptotic human DNA in stool [abstract]." Gastroenterology **118**(A855).
- Ahlquist, D. A. and J. A. Gilbert (1996). "Stool markers for colorectal cancer screening: future considerations." Dig Dis **14**(3): 132-44.
- Ahlquist, D. A., J. J. Harrington, et al. (2000). "Morphometric analysis of the "mucocellular layer" overlying colorectal cancer and normal mucosa: relevance to exfoliation and stool screening." Hum Pathol **31**(1): 51-7.
- Ahlquist, D. A., D. B. McGill, et al. (1989). "Patterns of occult bleeding in asymptomatic colorectal cancer." Cancer **63**(9): 1826-30.
- Ahlquist, D. A., J. E. Skoletsky, et al. (2000). "Colorectal cancer screening by detection of altered human DNA in stool: feasibility of a multitarget assay panel." Gastroenterology **119**(5): 1219-27.
- Ajioka, Y., L. J. Allison, et al. (1996). "Significance of MUC1 and MUC2 mucin expression in colorectal cancer." J Clin Pathol **49**(7): 560-4.
- Anderson, N. L. (1995). "Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it." Biotech.Gen.Eng. Rev. **13**: 19-50.

- Anderson, N. L. and N. G. Anderson (2002). "The human plasma proteome: history, character, and diagnostic prospects." Mol Cell Proteomics **1**(11): 845-67.
- Anker, N., F. Christensen, et al. (1974). "[Hemoccult. A new method for demonstration of occult bleeding]." Ugeskr Laeger **136**(36): 2025-7.
- Atkin, W. (2003). "Options for screening for colorectal cancer." Scand J Gastroenterol Suppl(237): 13-6.
- Atkin, W. S., J. Cuzick, et al. (1993). "Prevention of colorectal cancer by once-only sigmoidoscopy." Lancet **341**(8847): 736-40.
- Atkin, W. S., R. Edwards, et al. "Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial." Lancet **375**(9726): 1624-33.
- Baak, J. P., F. R. Path, et al. (2003). "Genomics and proteomics in cancer." Eur J Cancer **39**(9): 1199-215.
- Baba, S. (1997). "Recent advances in molecular genetics of colorectal cancer." World J Surg **21**(7): 678-87.
- Baggerly, K. A., J. S. Morris, et al. (2004). "Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments." Bioinformatics **20**(5): 777-85.
- Banks, R. E., M. J. Dunn, et al. (1999). "The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis--preliminary findings." Electrophoresis **20**(4-5): 689-700.
- Belanger, A., H. van Halbeek, et al. (1995). "Molecular mass and carbohydrate structure of prostate specific antigen: studies for establishment of an international PSA standard." Prostate **27**(4): 187-97.

- Berry, D. P., P. Clarke, et al. (1997). "Randomized trial of the addition of flexible sigmoidoscopy to faecal occult blood testing for colorectal neoplasia population screening." Br J Surg **84**(9): 1274-6.
- Bhattacharya, I. and E. M. Sack (1996). "Screening colonoscopy: the cost of common sense." Lancet **347**(9017): 1744-5.
- Blanks, R. G., S. M. Moss, et al. (2000). "Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality." Bmj **321**(7262): 665-9.
- Bowles, C. J., R. Leicester, et al. (2004). "A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow?" Gut **53**(2): 277-83.
- Brenner, D. J. and M. A. Georgsson (2005). "Mass screening with CT colonography: should the radiation exposure be of concern?" Gastroenterology **129**(1): 328-37.
- Brouwer, J. and F. Smekens (1991). "Instability of fecal lysozyme." Clin Chim Acta **201**(1-2): 137-8.
- Brown, G. and I. R. Daniels (2005). "Preoperative staging of rectal cancer: the MERCURY research project." Recent Results Cancer Res **165**: 58-74.
- Buderus, S., J. Boone, et al. (2004). "Fecal lactoferrin: a new parameter to monitor infliximab therapy." Dig Dis Sci **49**(6): 1036-9.
- Byrd, J. C. and R. S. Bresalier (2004). "Mucins and mucin binding proteins in colorectal cancer." Cancer Metastasis Rev **23**(1-2): 77-99.
- Cairns, S. and J. H. Scholefield (2002). "Guidelines for colorectal cancer screening in high risk groups." Gut **51 Suppl 5**: V1-2.

- Calvert, P. M. and H. Frucht (2002). "The genetics of colorectal cancer." Ann Intern Med **137**(7): 603-12.
- Carter, H. B. and W. B. Isaacs (2004). "Improved biomarkers for prostate cancer: a definite need." J Natl Cancer Inst **96**(11): 813-5.
- Celis, J. E. and P. Gromov (2003). "Proteomics in translational cancer research: toward an integrated approach." Cancer Cell **3**(1): 9-15.
- Celis, J. E., H. Wolf, et al. (2000). "Bladder squamous cell carcinoma biomarkers derived from proteomics." Electrophoresis **21**(11): 2115-21.
- Chen, C. C., S. H. Yang, et al. (2005). "Is it reasonable to add preoperative serum level of CEA and CA19-9 to staging for colorectal cancer?" J Surg Res **124**(2): 169-74.
- Cole, S. R. and G. P. Young (2001). "Effect of dietary restriction on participation in faecal occult blood test screening for colorectal cancer." Med J Aust **175**(4): 195-8.
- Coleman, M. P. (2004). "Trends and socioeconomic inequalities in cancer survival in England and Wales up to 2001." Br J Cancer **90**((7)): 1367-73.
- Congress, B. S. S. f. t. t. (2002). "Expressing the sense of the Congress that the Secretary of Health and Human Services should conduct or support research on certain tests to screen for ovarian cancer, and Federal health care programs and group and individual health plans should cover the tests if demonstrated to be effective, and for other purposes." US Government H.CON.RES.385: Rep Israel, Steve Rep DeLauro, Rosa L.
- Conlin, A., G. Smith, et al. (2005). "The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma." Gut **54**(9): 1283-6.
- Conrads, T. P., M. Zhou, et al. (2003). "Cancer diagnosis using proteomic patterns." Expert Rev Mol Diagn **3**(4): 411-20.

- Coombes, K. R., H. A. Fritsche, Jr., et al. (2003). "Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization." Clin Chem **49**(10): 1615-23.
- Dachman, A. H. and H. Yoshida (2003). "Virtual colonoscopy: past, present, and future." Radiol Clin North Am **41**(2): 377-93.
- Dave, S., S. Hui, et al. (2003). "Is the distal hyperplastic polyp a marker for proximal neoplasia?" J Gen Intern Med **18**(2): 128-37.
- Davies, R. J., A. Freeman, et al. (2002). "Analysis of minichromosome maintenance proteins as a novel method for detection of colorectal cancer in stool." Lancet **359**(9321): 1917-9.
- Diamandis, E. P. (2003). "Point: Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics?" Clin Chem **49**(8): 1272-5.
- Digby, J., C. G. Fraser, et al. "Faecal haemoglobin concentration is related to severity of colorectal neoplasia." J Clin Pathol **66**(5): 415-9.
- Doran, J. and J. D. Hardcastle (1982). "Bleeding patterns in colorectal cancer: the effect of aspirin and the implications for faecal occult blood testing." Br J Surg **69**(12): 711-3.
- Dubrow, R., C. S. Kim, et al. (1992). "Fecal lysozyme: an unreliable marker for colorectal cancer." Am J Gastroenterol **87**(5): 617-21.
- Duffy, M. J., A. van Dalen, et al. (2003). "Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines." European Journal of Cancer **39**(6): 718-727.
- Dukes, C. (1932). "The classification of cancer of the rectum." The Journal of Pathology and Bacteriology **35**: 323-32.

- Dulucq, J. L., P. Wintringer, et al. (2005). "Laparoscopic rectal resection with anal sphincter preservation for rectal cancer: long-term outcome." Surg Endosc **19**(11): 1468-74.
- Eastwood, G. L. (1977). "Gastrointestinal epithelial renewal." Gastroenterology **72**(5 Pt 1): 962-75.
- Eastwood, G. L. (1992). "Epithelial renewal in premalignant conditions of the gastrointestinal tract: a review." J Clin Gastroenterol **14 Suppl 1**: S29-33.
- Eastwood, G. L. (1995). "A review of gastrointestinal epithelial renewal and its relevance to the development of adenocarcinomas of the gastrointestinal tract." J Clin Gastroenterol **21 Suppl 1**: S1-11.
- Eide, T. J. (1991). "Natural history of adenomas." World J Surg **15**(1): 3-6.
- Executive, S. (2007). Colorectal Cancer Keypoints. WebSite. **2009**:
- Fearon, E. R. and B. Vogelstein (1990). "A genetic model for colorectal tumorigenesis." Cell **61**(5): 759-67.
- Feldman, A. L., V. Espina, et al. (2004). "Use of proteomic patterns to screen for gastrointestinal malignancies." Surgery **135**(3): 243-7.
- Fidler, J. L., C. D. Johnson, et al. (2002). "Detection of flat lesions in the colon with CT colonography." Abdom Imaging **27**(3): 292-300.
- Finan, P. J., J. K. Ritchie, et al. (1987). "Synchronous and 'early' metachronous carcinomas of the colon and rectum." Br J Surg **74**(10): 945-7.
- Fletcher, R. H. (1986). "Carcinoembryonic antigen." Ann Intern Med **104**(1): 66-73.
- Fraser, C. G., C. M. Matthew, et al. (2006). "Immunochemical testing of individuals positive for guaiac faecal occult blood test in a screening programme for colorectal cancer: an observational study." Lancet Oncol **7**(2): 127-31.

- Frazier, A. L., G. A. Colditz, et al. (2000). "Cost-effectiveness of screening for colorectal cancer in the general population." Jama **284**(15): 1954-61.
- Fuszek, P., P. Lakatos, et al. (2004). "[Significance of serum calcium levels in colorectal cancer]." Orv Hetil **145**(30): 1561-5.
- Fuszek, P., P. Lakatos, et al. (2004). "Relationship between serum calcium and CA 19-9 levels in colorectal cancer." World J Gastroenterol **10**(13): 1890-2.
- Gatta, G., J. Faivre, et al. (1998). "Survival of colorectal cancer patients in Europe during the period 1978-1989." Eur J Cancer **34**(14): 2176-2183.
- Ghnassia, J. P., J. F. Rodier, et al. (2001). "Elevation of CA15-3 caused by a benign tumour." Lancet Oncol **2**(2): 102.
- Gilbert, J. A., D. A. Ahlquist, et al. (1996). "Fecal marker variability in colorectal cancer: calprotectin versus hemoglobin." Scand J Gastroenterol **31**(10): 1001-5.
- Gluecker, T. M., C. D. Johnson, et al. (2003). "Extracolonic findings at CT colonography: evaluation of prevalence and cost in a screening population." Gastroenterology **124**(4): 911-6.
- Greiner, K. A., K. K. Engelman, et al. (2004). "Barriers to colorectal cancer screening in rural primary care." Prev Med **38**(3): 269-75.
- Gretzer, M. B., D. W. Chan, et al. (2004). "Proteomic analysis of dunning prostate cancer cell lines with variable metastatic potential using SELDI-TOF." Prostate **60**(4): 325-31.
- Griffiths, E. K. and D. V. Schapira (1991). "Serum ferritin and stool occult blood and colon cancer screening." Cancer Detect Prev **15**(4): 303-5.
- Group, U. C. C. S. P. (2004). "Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom." Bmj **329**(7458): 133.

- Gyde, S. (1990). "Screening for colorectal cancer in ulcerative colitis: dubious benefits and high costs." Gut **31**(10): 1089-92.
- Hammarstrom, S. (1999). "The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues." Semin Cancer Biol **9**(2): 67-81.
- Hara, A. K., C. D. Johnson, et al. (2000). "Incidental extracolonic findings at CT colonography." Radiology **215**(2): 353-7.
- Hardcastle, J. D., J. O. Chamberlain, et al. (1996). "Randomised controlled trial of faecal-occult-blood screening for colorectal cancer." Lancet **348**(9040): 1472-7.
- Hardt, P. D., M. Toepler, et al. (2003). "Measurement of fecal pyruvate kinase type M2 (tumor M2-PK) concentrations in patients with gastric cancer, colorectal cancer, colorectal adenomas and controls." Anticancer Res **23**(2A): 851-3.
- Hardy, R. G., S. J. Meltzer, et al. (2000). "ABC of colorectal cancer. Molecular basis for risk factors." Bmj **321**(7265): 886-9.
- Heald, R. J. and R. D. Ryall (1986). "Recurrence and survival after total mesorectal excision for rectal cancer." Lancet **1**(8496): 1479-82.
- Hewitson, P., P. Glasziou, et al. (2008). "Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update." Am J Gastroenterol **103**(6): 1541-9.
- Hlavaty, J. J., Alan W. Partin, Felicity Kusnitz, Matthew J. Shue, Adam Stieg, Kate Bennett and Joseph V. Briggman (2001). "Mass Spectroscopy as a Discovery Tool for Identifying Serum Markers for Prostate Cancer." Clin Chem **47**(November): 1924-1926.

- Hlavaty, J. J., A. W. Partin, et al. (2003). "Identification and preliminary clinical evaluation of a 50.8-kDa serum marker for prostate cancer." Urology **61**(6): 1261-5.
- Hoff, G., T. Grotmol, et al. (2004). "Testing for faecal calprotectin (PhiCal) in the Norwegian Colorectal Cancer Prevention trial on flexible sigmoidoscopy screening: comparison with an immunochemical test for occult blood (FlexSure OBT)." Gut **53**(9): 1329-33.
- Homa, K., M. Brzosko, et al. (2005). "[Acceptance of screening colonoscopies in the prevention of colorectal cancer in relation to some demographic factors]." Pol Merkurusz Lek **19**(110): 179-82.
- Imperiale, T. F., D. R. Wagner, et al. (2000). "Risk of Advanced Proximal Neoplasms in Asymptomatic Adults According to the Distal Colorectal Findings." N Engl J Med **343**(3): 169-174.
- Imperiale, T. F., D. R. Wagner, et al. (2002). "Results of screening colonoscopy among persons 40 to 49 years of age." N Engl J Med **346**(23): 1781-5.
- Johnson, C. D., W. S. Harmsen, et al. (2003). "Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps." Gastroenterology **125**(2): 311-9.
- Kang, H. J., H. M. Chang, et al. (2005). "Phase II study of capecitabine and cisplatin as first-line combination therapy in patients with gastric cancer recurrent after fluoropyrimidine-based adjuvant chemotherapy." Br J Cancer **92**(2): 246-51.
- Kemp, Z., C. Thirlwell, et al. (2004). "An update on the genetics of colorectal cancer." Hum Mol Genet **13 Spec No 2**: R177-85.

- Kim, S. B., L. C. Fernandes, et al. (2003). "Assessment of the value of preoperative serum levels of CA 242 and CEA in the staging and postoperative survival of colorectal adenocarcinoma patients." Int J Biol Markers **18**(3): 182-7.
- Koopmann, J., Z. Zhang, et al. (2004). "Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry." Clin Cancer Res **10**(3): 860-8.
- Kristinsson, J., C. H. Armbruster, et al. (2001). "Fecal excretion of calprotectin in colorectal cancer: relationship to tumor characteristics." Scand J Gastroenterol **36**(2): 202-7.
- Kronborg, O., C. Fenger, et al. (1996). "Randomised study of screening for colorectal cancer with faecal-occult-blood test." Lancet **348**(9040): 1467-71.
- Lacy, A. M., J. C. Garcia-Valdecasas, et al. (1995). "Short-term outcome analysis of a randomized study comparing laparoscopic vs open colectomy for colon cancer." Surg Endosc **9**(10): 1101-5.
- Leslie, A., F. A. Carey, et al. (2002). "The colorectal adenoma-carcinoma sequence." Br J Surg **89**(7): 845-60.
- Levi, F., L. Randimbison, et al. (1993). "Trends in subsite distribution of colorectal cancers and polyps from the Vaud Cancer Registry." Cancer **72**(1): 46-50.
- Levin, B., D. Brooks, et al. (2003). "Emerging technologies in screening for colorectal cancer: CT colonography, immunochemical fecal occult blood tests, and stool screening using molecular markers." CA Cancer J Clin **53**(1): 44-55.
- Lewis, J. D., K. Ng, et al. (2003). "Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy." Arch Intern Med **163**(4): 413-20.

- Li, J., Z. Zhang, et al. (2002). "Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer." Clin Chem **48**(8): 1296-304.
- Libby, G., D. H. Brewster, et al. "The impact of population-based faecal occult blood test screening on colorectal cancer mortality: a matched cohort study." Br J Cancer **107**(2): 255-9.
- Limburg, P. J., D. A. Ahlquist, et al. (2000). "Immunodiscrimination of colorectal neoplasia using MUC1 antibodies: discrepant findings in tissue versus stool." Dig Dis Sci **45**(3): 494-9.
- Limburg, P. J., M. E. Devens, et al. (2003). "Prospective evaluation of fecal calprotectin as a screening biomarker for colorectal neoplasia." Am J Gastroenterol **98**(10): 2299-305.
- Lin, O. S., D. B. Schembre, et al. (2005). "Risk of proximal colorectal neoplasia among asymptomatic patients with distal hyperplastic polyps." The American Journal of Medicine **118**(10): 1113-1119.
- Lindholm, E., H. Brevinge, et al. (2008). "Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer." Br J Surg **95**(8): 1029-36.
- Lindmark, G., R. Bergstrom, et al. (1995). "The association of preoperative serum tumour markers with Dukes' stage and survival in colorectal cancer." Br J Cancer **71**(5): 1090-4.
- Llor, X., E. Pons, et al. (2005). "Differential features of colorectal cancers fulfilling Amsterdam criteria without involvement of the mutator pathway." Clin Cancer Res **11**(20): 7304-10.
- Louhimo, J., H. Alfthan, et al. (2004). "Serum HCG beta and CA 72-4 are stronger prognostic factors than CEA, CA 19-9 and CA 242 in pancreatic cancer." Oncology **66**(2): 126-31.

- Macafee, D. A. and J. H. Scholefield (2003). "Antagonist: population based endoscopic screening for colorectal cancer." Gut **52**(3): 323-6.
- Mandel, J. S., J. H. Bond, et al. (1993). "Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study." N Engl J Med **328**(19): 1365-71.
- Mandel, J. S., T. R. Church, et al. (2000). "The effect of fecal occult-blood screening on the incidence of colorectal cancer." N Engl J Med **343**(22): 1603-7.
- Mayor, S. (1998). "Cancer is main cause of death in Britain." BMJ **316**(7131): 571.
- McArdle, C. (2000). "ABC of colorectal cancer: Effectiveness of follow up." BMJ **321**(7272): 1332-1335.
- McKerrow, J. H., V. Bhargava, et al. (2000). "A functional proteomics screen of proteases in colorectal carcinoma." Mol Med **6**(5): 450-60.
- Melle, C., G. Ernst, et al. (2005). "Discovery and identification of alpha-defensins as low abundant, tumor-derived serum markers in colorectal cancer." Gastroenterology **129**(1): 66-73.
- Morita, S., T. Nomura, et al. (2004). "Does serum CA19-9 play a practical role in the management of patients with colorectal cancer?" Dis Colon Rectum **47**(2): 227-32.
- Nakagoe, T., T. Sawai, et al. (2003). "Preoperative serum level of CA19-9 predicts recurrence after curative surgery in node-negative colorectal cancer patients." Hepatogastroenterology **50**(51): 696-9.
- Network, S. I. G. (2011). "Management of Colorectal Cancer."
- Nicholson, F. B., J. L. Barro, et al. (2005). "The role of CT colonography in colorectal cancer screening." Am J Gastroenterol **100**(10): 2315-23.

- Nishizuka, S., S. T. Chen, et al. (2003). "Diagnostic markers that distinguish colon and ovarian adenocarcinomas: identification by genomic, proteomic, and tissue array profiling." Cancer Res **63**(17): 5243-50.
- Ornstein, D. K., W. Rayford, et al. (2004). "Serum proteomic profiling can discriminate prostate cancer from benign prostates in men with total prostate specific antigen levels between 2.5 and 15.0 ng/ml." J Urol **172**(4 Pt 1): 1302-5.
- Parsi, M. A., B. Shen, et al. (2004). "Fecal lactoferrin for diagnosis of symptomatic patients with ileal pouch-anal anastomosis." Gastroenterology **126**(5): 1280-6.
- Patterson, S. D. (2003). "Proteomics: the first decade and beyond." Nature **33**: 311-323.
- Petricoin, E. E., C. P. Paweletz, et al. (2002). "Clinical applications of proteomics: proteomic pattern diagnostics." J Mammary Gland Biol Neoplasia **7**(4): 433-40.
- Petricoin, E. F., 3rd, D. K. Ornstein, et al. (2002). "Serum proteomic patterns for detection of prostate cancer." J Natl Cancer Inst **94**(20): 1576-8.
- Petricoin, E. F., A. M. Ardekani, et al. (2002). "Use of proteomic patterns in serum to identify ovarian cancer." Lancet **359**(9306): 572-7.
- Petricoin, E. F. and L. A. Liotta (2004). "SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer." Curr Opin Biotechnol **15**(1): 24-30.
- Podolsky, D. K. (2000). "Going the distance--the case for true colorectal-cancer screening." N Engl J Med **343**(3): 207-8.
- Pusztai, L., B. W. Gregory, et al. (2004). "Pharmacoproteomic analysis of prechemotherapy and postchemotherapy plasma samples from patients receiving neoadjuvant or adjuvant chemotherapy for breast carcinoma." Cancer **100**(9): 1814-22.

- Qu, Y., B. L. Adam, et al. (2002). "Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients." Clin Chem **48**(10): 1835-43.
- Rabilloud, T. (2002). "Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains." Proteomics **2**(1): 3-10.
- Rasmussen, M., O. Kronborg, et al. (1999). "Possible advantages and drawbacks of adding flexible sigmoidoscopy to hemoccult-II in screening for colorectal cancer. A randomized study." Scand J Gastroenterol **34**(1): 73-8.
- Rex, D. K., Cummings OW, Helper DJ et al. (1996). "5-year incidence of adenomas after negative colonoscopy in asymptomatic average-risk persons." Gastroenterology **112**: 1178-1181`.
- Rex, D. K., D. A. Johnson, et al. (2000). "Colorectal cancer prevention 2000: screening recommendations of the American College of Gastroenterology. American College of Gastroenterology." Am J Gastroenterol **95**(4): 868-77.
- Rodland, K. D. (2004). "Proteomics and cancer diagnosis: the potential of mass spectrometry." Clin Biochem **37**(7): 579-83.
- Roessler, M., W. Rollinger, et al. (2005). "Identification of nicotinamide N-methyltransferase as a novel serum tumor marker for colorectal cancer." Clin Cancer Res **11**(18): 6550-7.
- Rogers, M. A., P. Clarke, et al. (2003). "Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility." Cancer Res **63**(20): 6971-83.

- Rogers MA, C. P., Noble J et al. (2003). "Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility." Cancer Res **63**: 6971-83.
- Rosenfeld, L. (1987). "Henry Bence Jones (1813-1873): the best "chemical doctor" in London." Clin Chem **33**(9): 1687-1692.
- Roseth, A. G., M. K. Fagerhol, et al. (1992). "Assessment of the neutrophil dominating protein calprotectin in feces. A methodologic study." Scand J Gastroenterol **27**(9): 793-8.
- Roseth, A. G., J. Kristinsson, et al. (1993). "Faecal calprotectin: a novel test for the diagnosis of colorectal cancer?" Scand J Gastroenterol **28**(12): 1073-6.
- Rubio, C. A. (2003). "Colorectal adenomas produce lysozyme." Anticancer Res **23**(6D): 5165-71.
- Rui, Z., J. Jian-Guo, et al. (2003). "Use of serological proteomic methods to find biomarkers associated with breast cancer." Proteomics **3**(4): 433-9.
- Ruiz-Pelaez, J. G. and S. Mattar (1999). "Accuracy of fecal lactoferrin and other stool tests for diagnosis of invasive diarrhea at a Colombian pediatric hospital." Pediatr Infect Dis J **18**(4): 342-6.
- Rustgi, A. K. (1993). "Molecular genetics and colorectal cancer." Gastroenterology **104**(4): 1223-5.
- Saitoh, O., K. Kojima, et al. (2000). "Comparison of tests for fecal lactoferrin and fecal occult blood for colorectal diseases: a prospective pilot study." Intern Med **39**(10): 778-82.

- Sauter, E. R., S. Shan, et al. (2005). "Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS." Int J Cancer **114**(5): 791-6.
- Sauter, E. R., W. Zhu, et al. (2002). "Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer." Br J Cancer **86**(9): 1440-3.
- Scerpella, E. G., P. C. Okhuysen, et al. (1994). "Evaluation of a New Latex Agglutination Test for Fecal Lactoferrin in Travelers' Diarrhea." J Travel Med **1**(2): 68-71.
- Scholefield, J. H. (2000). "ABC of colorectal cancer: screening." Bmj **321**(7267): 1004-6.
- Scholefield, J. H., S. Moss, et al. (2002). "Effect of faecal occult blood screening on mortality from colorectal cancer: results from a randomised controlled trial." Gut **50**(6): 840-4.
- Scholefield, J. H., M. H. Robinson, et al. (1998). "Serum ferritin. Screening test for colorectal cancer?" Dis Colon Rectum **41**(8): 1029-31; discussion 1031-2.
- Segnan, N., C. Senore, et al. (2005). "Randomized Trial of Different Screening Strategies for Colorectal Cancer: Patient Response and Detection Rates
10.1093/jnci/dji050." J Natl Cancer Inst **97**(5): 347-357.
- Selby, J. V., G. D. Friedman, et al. (1992). "A case-control study of screening sigmoidoscopy and mortality from colorectal cancer." N Engl J Med **326**(10): 653-7.
- Shiwa, M., Y. Nishimura, et al. (2003). "Rapid discovery and identification of a tissue-specific tumor biomarker from 39 human cancer cell lines using the SELDI ProteinChip platform." Biochem Biophys Res Commun **309**(1): 18-25.
- Sidransky, D., T. Tokino, et al. (1992). "Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors." Science **256**(5053): 102-5.
- Slater, G., A. E. Papatestas, et al. (1982). "Age distribution of right- and left-sided colorectal cancers." Am J Gastroenterol **77**(2): 63-6.

- Soltys, S. (2004). "The Use of Plasma Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Proteomic Patterns for Detection of Head and Neck Squamous Cell Cancers." Clinical Cancer Research **10**: 4806-4812.
- Song, S., J. C. Byrd, et al. (2005). "Galectin-3 Modulates MUC2 Mucin Expression in Human Colon Cancer Cells at the Level of Transcription via AP-1 Activation." Gastroenterology **129**(5): 1581-91.
- Sorace, J. M. and M. Zhan (2003). "A data review and re-assessment of ovarian cancer serum proteomic profiling." BMC Bioinformatics **4**(1): 24.
- Statistics, S. H. (2005). "<http://www.isdscotland.org/isd/>."
- Steele, R. J., P. L. McClements, et al. (2009). "Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer." Gut **58**(4): 530-5.
- Steiner, T. S., C. A. Flores, et al. (1997). "Fecal lactoferrin, interleukin-1beta, and interleukin-8 are elevated in patients with severe Clostridium difficile colitis." Clin Diagn Lab Immunol **4**(6): 719-22.
- Sudo, I., M. Igawa, et al. (1993). "[A study to determine fecal lactoferrin in patients with ulcerative colitis]." Nippon Shokakibyo Gakkai Zasshi **90**(4): 824.
- Sugi, K., O. Saitoh, et al. (1996). "Fecal lactoferrin as a marker for disease activity in inflammatory bowel disease: comparison with other neutrophil-derived proteins." Am J Gastroenterol **91**(5): 927-34.
- Sung, J. J., J. Y. Lau, et al. (2005). "Increasing incidence of colorectal cancer in Asia: implications for screening." Lancet Oncol **6**(11): 871-6.
- Tibble, J., G. Sigthorsson, et al. (2001). "Faecal calprotectin and faecal occult blood tests in the diagnosis of colorectal carcinoma and adenoma." Gut **49**(3): 402-8.

- Tolson, J., R. Bogumil, et al. (2004). "Serum protein profiling by SELDI mass spectrometry: detection of multiple variants of serum amyloid alpha in renal cancer patients." Lab Invest **84**(7): 845-56.
- Towler, B., L. Irwig, et al. (1998). "A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemocult." Bmj **317**(7158): 559-65.
- UK, C. R. (2006). UK cancer mortality statistics for common cancers.
- UK Flexible Sigmoidoscopy Screening Trial Investigators "Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial." Lancet **359**(9314): 1291-300.
- UK Colorectal Cancer Screening Pilot. 2004 "Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom." Bmj **329**(7458): 133.
- Verrills, N. M., B. J. Walsh, et al. (2003). "Proteome analysis of vinca alkaloid response and resistance in acute lymphoblastic leukemia reveals novel cytoskeletal alterations." J Biol Chem **278**(46): 45082-93.
- Villa, E., A. Dugani, et al. (1996). "Identification of subjects at risk for colorectal carcinoma through a test based on K-ras determination in the stool." Gastroenterology **110**(5): 1346-53.
- Walsh, J. M. and J. P. Terdiman (2003). "Colorectal cancer screening: clinical applications." Jama **289**(10): 1297-302.
- Walsh, J. M. and J. P. Terdiman (2003). "Colorectal cancer screening: scientific review." Jama **289**(10): 1288-96.
- Walter, L. C., C. L. Lewis, et al. (2005). "Screening for colorectal, breast, and cervical cancer in the elderly: a review of the evidence." Am J Med **118**(10): 1078-86.

- Warren, J. D., W. Xiong, et al. "Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer." BMC Med **9**: 133.
- Weitz, J. x. f. r., M. Koch, et al. "Colorectal cancer." The Lancet **365**(9454): 153-165.
- Whynes, D. K. (2004). "Cost-effectiveness of screening for colorectal cancer: evidence from the Nottingham faecal occult blood trial." J Med Screen **11**(1): 11-5.
- Wilson JM, J. F. (1968). "Principles and practice of screening for disease." Public Health Papers **No.34**(Geneva).
- Winawer, S. J., B. J. Flehinger, et al. (1993). "Screening for colorectal cancer with fecal occult blood testing and sigmoidoscopy." J Natl Cancer Inst **85**(16): 1311-8.
- Wong, Y. F., T. H. Cheung, et al. (2004). "Protein profiling of cervical cancer by protein-biochips: proteomic scoring to discriminate cervical cancer from normal cervix." Cancer Lett **211**(2): 227-34.
- Woodman, C. B., P. Prior, et al. (1995). "Prospects for the secondary prevention of colorectal cancer: screening by flexible sigmoidoscopy?" J Med Screen **2**(2): 71-8.
- Yu, J. K., Y. D. Chen, et al. (2004). "An integrated approach to the detection of colorectal cancer utilizing proteomics and bioinformatics." World J Gastroenterol **10**(21): 3127-31.
- aubert, A. G. (1997). "Can surveillance intervals be lengthened following colonoscopic polypectomy(abstract)." Gastroenterology(112): A50.
- Zhang, W. and B. T. Chait (2000). "ProFound: an expert system for protein identification using mass spectrometric peptide mapping information." Anal Chem **72**(11): 2482-9.
- Zhao, G., C. F. Gao, et al. (2004). "[Identification of colorectal cancer using proteomic patterns in serum]." Ai Zheng **23**(6): 614-8.
- Zheng, S. (2004). "Personal correspondance with author."

Zheng, Y., Y. Xu, et al. (2003). "Prostate carcinoma tissue proteomics for biomarker discovery." Cancer **98**(12): 2576-82.